

Result No.	Score	Query Match	Length	DB ID	Description
1	60	96.8	10	23 ABG32532	zinc finger protein
2	47	75.8	67	23 ABP2773	Human ovarian anti
3	46	74.2	638	22 ABG07106	Novel human diagno
4	46	74.2	638	22 ABG08897	Novel human diagno
5	46	74.2	638	22 ABG24363	Novel human diagno
6	45.5	73.4	84	21 AAY98345	Synthetic protein
7	45.5	73.4	90	23 ABP48203	Sp-1 zinc finger p
8	45.5	73.4	100	21 ABA14294	Transcription fact
9	45.5	73.4	100	23 ABJ03954	zinc finger protein

31-JAN-2001; 2001US-265782P.

31-JAN-2001; 2001US-265782P.

PR 31-JAN-2001; 2001US-265858P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Coleman TA, Mansfield B;
 PI WPI; 2002-643357/69.
 XX
 XX Novel scaffolded fusion polypeptide useful for therapeutic purposes or
 XX for screening molecules that bind/activate/inhibit/modulate the
 XX polypeptide, comprises a functional polypeptide domain fused to a
 XX scaffold domain
 XX
 XX Claim 5; Page 16; 64pp; English.
 XX
 XX The invention relates to a scaffolded fusion polypeptide comprising a
 XX functional polypeptide domain fused to a scaffold domain, where the
 XX functional polypeptide domain corresponds to a soluble loop of an
 XX integral membrane protein (e.g. human CCR5, a transmembrane receptor
 XX involved in HIV (human immunodeficiency virus) infection).
 XX Also included are: (1) a polypeptide comprising a scaffold domain; (2) a
 XX nucleic acid encoding the fusion polypeptide; (3) a vector cassette for
 XX the expression of the fusion polypeptide comprising an expression region
 XX operably linked to a promoter, where the expression region comprises a
 XX number of cassettes, each of which encodes a module, domain or strand of
 XX the fusion polypeptide and (4) a host cell comprising the vector or
 XX nucleic acid. The fusion polypeptide is useful for screening molecules
 XX that bind/activate/inhibit/modulate the fusion polypeptide, by expressing
 XX the fusion polypeptide from and identifying a molecule that binds to the
 XX fusion polypeptide. The fusion polypeptide is useful in diagnostic
 XX methods, in assays to identify compounds that interact with loops of
 XX fragments of an extracellular domain (ECD) or an intracellular domain
 XX (ICD) or to rapidly assay the function of mutated portions of mutant
 XX integral membrane proteins without having to produce significant
 XX quantities of the entire mutant integral membrane protein, to generate
 XX antibodies that recognise the integral membrane proteins from which they
 XX are designed, to competitively bind the ligand of a naturally occurring
 XX receptor in vitro or in vivo, to display and/or screen soluble domains
 XX from protein such as integral membrane proteins, to probe the structure
 XX of ECD or ICD, or both, of an integral protein membrane, to modulate the
 XX activity of a receptor in vivo, and for treating or preventing viral
 XX infection, preferably human HIV infection e.g. by gene therapy using the
 XX encoding nucleic acid. The present sequence is a Zinc finger protein of
 XX metal chelating motif suitable for inclusion in the fusion protein of the
 XX invention.
 XX
 XX Sequence 10 AA;
 XX
 XX Query Match 96.8%; Score 60; DB 23; Length 10;
 XX Best Local Similarity 100.0%; Pred. No. 0.0013;
 XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 CPECKXHQTH 10
 XX | | | | | | | | | |
 XX Db 1 CPECKXHQTH 10
 XX
 XX RESULT 2
 XX ABP42773
 XX ID ABP42773 standard; Protein: 67 AA.
 XX AC ABP42773;
 XX
 XX DT 22-AUG-2002 (first entry)
 XX
 XX DE Human ovarian antigen HPAMC04, SEQ ID NO:3905.
 XX
 XX Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
 XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 XX inflammatory condition; immune disorder; blood disorder;
 XX cardiovascular disorder; respiratory disorder; neurological disorder;

KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX
 XX Homo sapiens.
 XX
 XX WO200200677-A1.
 XX
 XX 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US18569.
 XX
 XX 07-JUN-2000; 2000US-209467P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 XX
 XX WPI; 2002-147878/19.
 XX
 XX N-PSDB; ABQ55850.
 XX
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 XX useful in the prevention, treatment and diagnosis of cancer (e.g.
 XX ovarian cancer), immune disorders, cardiovascular disorders and
 XX neurological diseases -
 XX
 XX Claim 11; SEQ ID NO 3905; 2922pp; English.
 XX
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 XX encompasses polypeptides 90% identical and polynucleotides 95% identical
 XX to the sequences of the invention. The invention additionally relates to
 XX recombinant vectors and host cells comprising human ovarian antigen
 XX polynucleotides, antibodies against human ovarian antigens, and the use
 XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
 XX treating, prognosing or preventing various ovary and/or breast-related
 XX disorders. Such conditions include ovarian cancer and breast cancer, and
 XX metastatic tumours of ovarian or breast origin, reproductive system
 XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
 XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 XX vaginitis), immune disorders (e.g., congenital and acquired
 XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
 XX respiratory disorders, neurological disorders, gastrointestinal disorders
 XX and urinary system disorders. Ovarian antigen polypeptides and
 XX polynucleotides may also be used in screening for compounds which
 XX modulate ovarian antigen expression or activity. The polynucleotides may
 XX further be used for gene therapy, chromosome mapping, in the
 XX identification of individuals and in forensic analysis, and the
 XX polypeptides may be used as food additives or to prepare antibodies
 XX useful in disease diagnosis, drug targeting and phenotyping. The present
 XX sequence represents a human ovarian antigen of the invention.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences.

Sequence 67 AA;

Query Match 75.8%; Score 47; DB 23; Length 67;

Best Local Similarity 70.0%; Pred. No. 1.1;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CPECKXHQTH 10

Db 49 CTECHQRLH 58

RESULT 3

ABG07106

ID ABG07106 standard; Protein: 638 AA.

XX	ABG07106;
DT	AC
XX	13-FEB-2002 (first entry)
DE	XX
XX	Novel human diagnostic protein #7097.
DE	XX
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder.
XX	Homo sapiens.
OS	WO200175067-A2.
PN	11-OCT-2001.
PD	XX
XX	30-MAR-2001; 2001WO-US08631.
PF	XX
XX	31-MAR-2000; 2000US-0540217.
PR	XX
XX	23-AUG-2000; 2000US-0649167.
PR	(HYSE-) HYSEQ INC.
PA	Drmanac RT, Liu C, Tang YT;
PI	WPI: 2001-639362/73.
PI	N-PSDB; AAS71293.
DR	New isolated polynucleotide and encoded polypeptides, useful in
DR	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
PT	Claim 20; SEQ ID No 37465; 103pp; English.
PS	The invention relates to isolated polynucleotide (I) and
PS	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	Sequence 638 AA;
SQ	Query Match 74.2%; Score 46; DB 22; Length 638;
	Best Local Similarity 60.0%; Pred. No. 13;
	Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY	1 CPECXHQRTH 10
	: :
Db	397 CKQCVRHERTH 406
	RESULT 4
ABG08897	ID
ABG08897 standard; Protein; 638 AA.	ID
XX	ABG08897;
XX	18-FEB-2002 (first entry)

XX	13-FEB-2002 (first entry)
XX	Novel human diagnostic protein #8888.
DE	XX
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder.
XX	Homo sapiens.
OS	WO200175067-A2.
PN	11-OCT-2001.
PD	XX
XX	30-MAR-2001; 2001WO-US08631.
PF	XX
XX	31-MAR-2000; 2000US-0540217.
PR	XX
XX	23-AUG-2000; 2000US-0649167.
PR	(HYSE-) HYSEQ INC.
PA	Drmanac RT, Liu C, Tang YT;
PI	WPI: 2001-639362/73.
PI	N-PSDB; AAS73084.
DR	New isolated polynucleotide and encoded polypeptides, useful in
DR	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
PT	Claim 20; SEQ ID No 39256; 103pp; English.
PS	The invention relates to isolated polynucleotide (I) and
PS	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. ABG00010-ABG30377 represent novel human
CC	diagnostic amino acid sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	Sequence 638 AA;
SQ	Query Match 74.2%; Score 46; DB 22; Length 638;
	Best Local Similarity 60.0%; Pred. No. 13;
	Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY	1 CPECXHQRTH 10
	: :
Db	397 CKQCVRHERTH 406
	RESULT 5
ABG24363	ID
ABG24363 standard; Protein; 638 AA.	ID
XX	ABG24363;
XX	18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #24354.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR MPI: 2001-639362/73.
 XX DR N-PSDB; AAS88550.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX PS Claim 20; SEQ ID No 54722; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 638 AA;
 Query Match 74.2%; Score 46; DB 22; Length 638;
 Best Local Similarity 60.0%; Pred. No. 13;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CPECXHORTH 10
 I : I : I : I : I :
 Db 397 CKQCVHERTH 406

RESULT 6
 ID AAY88345
 ID AAY88345 standard; protein; 84 AA.

XX AC AAY88345;
 XX 14-JUL-2000 (first entry)
 XX Synthetic protein consisting of three zinc finger domains.

XX Zinc finger repeat sequence; gene library; specific binding interaction;
 KW protein identification method; selection; screening.
 XX OS Synthetic.
 XX PN WO200015777-A1.
 XX PD 23-MAR-2000.
 XX PF 14-SEP-1999; 99WO-GB03081.
 XX PR 14-SEP-1998; 98EP-0307434.
 XX (UYAS-) UNIV ASTON.
 XX (AMSH) AMERSHAM PHARMACIA BIOTECH UK LTD.
 XX PI Hine AV, Morgan LJ, Santos AF, Palfrey D;
 XX MPI: 2000-271418/23.
 XX Set of libraries of genes which code for proteins capable of specific
 PT binding interactions are used as selection or screening strategies in a
 PT combinatorial approach to obtain novel proteins capable of binding a
 PT desired target molecule.
 XX PS Example; Page 32; 48pp; English.
 XX This sequence represents a synthetic protein which contains three
 CC consensus zinc finger repeat sequences. The invention relates to a set of
 CC libraries of genes which code for proteins which are capable of specific
 CC binding interactions using amino acid residues at two or more determined
 CC positions. The present sequence represents a protein encoded by the genes
 CC in the libraries of the invention. The libraries of the invention can be
 CC used in a method of identifying a protein which interacts with a specific
 CC binding partner. The libraries and methods are useful as selection or
 CC screening strategies in a combinatorial approach to obtain novel proteins
 CC capable of binding a desired target molecule.
 XX SQ Sequence 84 AA;
 Query Match 73.4%; Score 45.5; DB 21; Length 84;
 Best Local Similarity 42.9%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 CPEC-----XHQRT 10
 I : I : I : I : I :
 Db 8 CPEGKFSKSHLVARQRT 28
 RESULT 7
 ID ABP48203 standard; Peptide; 90 AA.
 XX AC ABP48203;
 XX 28-AUG-2002 (first entry)
 XX Sp-1 zinc finger protein SEQ ID NO:15.
 XX Zinc finger protein; ZFP; DNA binding protein; zinc finger.
 XX OS Synthetic.
 XX PN WO200242459-A2.
 XX PD 30-MAY-2002.
 XX PF 20-NOV-2001; 2001WO-US43438.
 XX PR 20-NOV-2000; 2000US-0716637.
 XX (SANG-) SANGAMO BIOSCIENCES INC.

XX Liu Q;
 XX WPI; 2002-500284/53.
 XX New zinc finger protein that binds to target site, useful in studying
 PT gene function and for human therapeutics and plant engineering,
 PT comprises first, second and third zinc fingers, ordered from N- to
 PT C-terminus -
 XX Disclosure; Page 17; 8lpp; English.
 XX The present invention describes a zinc finger protein (I) that binds to
 CC a target site, comprising a first (F1), a second (F2), and a third (F3)
 CC zinc finger, ordered F1, F2, F3 from N-terminus to C-terminus, where the
 CC target site comprises, in 3'-5' direction, a first (S1), a second (S2),
 CC and a third (S3) target sub-site. Also described are: (I) a polypeptide
 CC (II) comprising (I); (2) a polynucleotide (III) encoding (I) or (II); and
 CC (3) designing (M) (I) involves selecting the F1 zinc finger such that
 CC it binds to the S1 target sub-site, selecting the F2 zinc finger such
 CC that it binds to the S2 target sub-site, and selecting the F3 zinc
 CC finger such that it binds to the S3 target sub-site, thus designing (I)
 CC that binds to a target site. (I) is useful for recognition of triplet
 CC target sub-sites having the nucleotide G in the 5'-most position of the
 CC sub-site. (I) is useful in studying gene function, and for human
 CC therapeutics and plant engineering. (I), (II) or (III) is useful in
 CC therapeutic methods to modulate the expression of a target region within
 CC a subject, in diagnostic methods for sequence specific detection of
 CC target nucleic acid in a sample, and in assays to determine the
 CC phenotype and function of gene expression. (I) has improved affinity
 CC and specificity for their target sequences, as well as enhanced
 CC biological activity. ABQ71213 to ABQ72214 and ABP48191 to ABP51230
 CC represent DNA target sequences and zinc finger peptides which are given
 CC in the exemplification of the present invention.
 XX Sequence 90 AA;
 SQ
 Query Match 73.4%; Score 45.5; DB 23; Length 90;
 Best Local Similarity 42.9%; Pred. No. 2.6;
 Matches 9; Conservative 0; Mismatches 1; Indels 11; Caps 1;
 QY 1 CPCC-----XHQTH 10
 DB 9 CPCCGKSFSSKSHLRAHQTH 29
 RESULT 8
 AAB14294
 ID AAB14294 standard; peptide; 100 AA.
 XX AAB14294;
 XX 08-MAR-2001 (first entry)
 XX Transcription factor Sp-1 protein fragment #2.
 DE Transcription factor: Sp-1; DNA binding domain; bacterial;
 KW viral; infection; oncogene; tumour suppressor gene; growth factor;
 KW cancer; CCR5 receptor; HIV-1; apoE; Alzheimer's disease;
 KW foetal haemoglobin; sickle cell anaemia; zinc finger protein.
 XX Unidentified.
 OS
 XX Key Location/Qualifiers
 XX Peptide 1..10
 FT /label= Leader_sequence
 FT
 XX WO200042219-A1.
 FN
 XX 20-JUL-2000.
 PD
 XX 06-JAN-2000; 2000WO-US00388.
 XX

PR 12-JAN-1999; 99US-0229007.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 XX Eisenberg SP, Case CC, Cox GN, Jamieson A, Rebar EJ;
 DR WPI; 2000-482840/42.
 XX Novel methods for selecting target sites for, and production of, zinc
 PI finger proteins, useful for controlling expression of target genes,
 PT e.g. for inhibiting oncogenes or treating sickle cell anemia -
 XX Disclosure; Page 16; 83pp; English.
 XX Zinc finger proteins (ZFPs) bind to DNA and can modulate the expression
 CC of a wide range of genes. The present invention relates to a method for
 CC selecting a target site within a nucleic acid to be targeted by a ZFP, by
 CC detecting a specific 10-base motif. The present sequence is the DNA
 CC binding domain of transcription factor Sp-1. This sequence is a ZFP. This
 CC sequence was used as a framework for designing or selecting a
 CC nonnaturally occurring or variant ZFP. ZFPs can be designed to bind to
 CC preselected target genes to modulate their expression i.e. activate or
 CC inhibit. The target genes may be involved in disease e.g. bacterial or
 CC viral genes (involved in bacterial or viral infections); oncogenes;
 CC tumour suppressor genes or growth factors (involved in cancer); CCR5
 CC receptor gene (implicated in HIV-1 infection) or apoE gene (implicated
 CC in Alzheimer's disease). In addition, foetal haemoglobin genes may be
 CC induced by a variant ZFP for treating sickle cell anaemia. Furthermore,
 CC plant genes involved in herbicide or disease resistance may be used as
 CC target genes, to improve plant characteristics.
 XX Sequence 100 AA;
 SQ
 Query Match 73.4%; Score 45.5; DB 21; Length 100;
 Best Local Similarity 42.9%; Pred. No. 2.9;
 Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
 QY 1 CPCC-----XHQTH 10
 DB 19 CPCCGKSFSSKSHLRAHQTH 39
 RESULT 9
 ABJ03954
 ID ABJ03954 standard; Peptide; 100 AA.
 XX ABJ03954;
 XX 25-SEP-2002 (first entry)
 DT Zinc finger protein Sp1 DNA binding domain SEQ ID NO: 223.
 DE Zinc finger protein; angiogenesis; vasculogenesis; ischaemia;
 KW diabetic retinopathy; psoriasis; arthropathy; cancer; tumour growth;
 KW gene therapy; antiatherosclerotic; vasotropic; antiarthritic; vulnarary;
 KW antituber; cytostatic; antiproliferative; antidiabetic; ophthalmological;
 KW osteopathic; antiinfertility.
 OS Unidentified.
 XX WO200246412-A2.
 PN
 XX 13-JUN-2002.
 PD
 XX 06-DEC-2001; 2001WO-US46861.
 PF
 XX 07-DEC-2000; 2000US-0733604.
 PR 12-DEC-2000; 2000US-0736083.
 PR 30-APR-2001; 2001US-0846033.
 XX (SANG-) SANGAMO BIOSCIENCES INC.
 PA Rebar E, Jamieson A, Liu Q, Liu P, Wolffe A, Eisenberg SP;
 PI

PI Jarvis E;
XX WPI: 2002-527918/56.
XX New zinc finger protein that binds to target site in vascular
PT endothelial growth factor gene, useful for modulating expression of the
PT gene and for treating atherosclerosis, ischemia, arthritis, wound or
PT ulcer
XX
PS Disclosure; Page 189; 195pp; English.
XX The present invention relates to a zinc finger protein that binds to a
CC target site in one or more vascular endothelial growth factor (VEGF)
CC genes. The protein is useful for modulating expression of a VEGF gene,
CC thereby regulating angiogenesis and vasculogenesis. This can be used to
CC treat atherosclerosis, ischemia, arthritis, wounds, ulcers, tumours,
CC diabetic retinopathy or psoriasis. The present sequence is a peptide
CC shown in the invention.
XX
XX Sequence 100 AA;
Query Match 73.4%; Score 45.5; DB 23; Length 100;
Best Local Similarity 42.9%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
QY 1 CPEC-----XHQTH 10
DB 19 CPEGKFSKSHLRAHQTH 39
RESULT 10
AAE30447
ID AAE30447 standard; Protein; 100 AA.
AC AAE30447;
XX
XX 24-FEB-2003 (first entry)
DE SPl transcription factor consensus protein fragment.
KW DNA binding specificity; zinc finger protein; specificity optimisation;
KW ZFP; gene expression; SPl transcription factor.
XX Unidentified.
XX
XX Key Location/Qualifiers
FH Peptide 1..10
FT /label= Signal_peptide
XX
PN WO20027227-A2.
PD
PD 03-OCT-2002.
XX
XX 20-NOV-2001; 2001WO-US43568.
XX
XX 20-NOV-2000; 2000US-0716637.
XX
XX (SANG-) SANGAMO BIOSCIENCES INC.
XX
XX Eisenberg SP, Liu Q, Jamieson A, Rebar E;
XX WPI: 2003-029936/02.
XX
XX Enhancing the binding specificity of a zinc finger protein, comprises
PT substituting amino acids in the protein for residues in the target
PT sequence to make a modified binding protein -
XX
XX Disclosure; Page 21; 55pp; English.
XX
XX The present invention relates to a novel method of enhancing the binding
CC specificity of a binding protein (zinc finger protein (ZFP)). The method
CC involves substituting one or more amino acids at positions in the binding
CC protein that affect the specificity of the binding protein for residues

CC In the target sequence to make a modified binding protein. The method is
CC useful in optimising the specificity of a binding protein, in modulating
CC the expression of a target gene in a subject and in diagnostic methods
CC for sequence-specific detection of a target nucleic acid in a sample.
CC The present sequence is SPl transcription factor consensus protein
CC fragment.
XX
XX Sequence 100 AA;
Query Match 73.4%; Score 45.5; DB 24; Length 100;
Best Local Similarity 42.9%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
QY 1 CPEC-----XHQTH 10
DB 19 CPEGKFSKSHLRAHQTH 39
RESULT 11
AAE06000
ID AAE06000 standard; Protein; 184 AA.
AC AAE06000;
XX
XX 24-SEP-2001 (first entry)
DE Zinc finger protein (ZFP) ml protein.
XX
XX Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture.
XX Unidentified.
XX
XX Key Location/Qualifiers
FH Misc-difference 23
FT /note= "Encoded by CAG"
FT Misc-difference 24
FT /note= "Encoded by AGC"
FT Misc-difference 25
FT /note= "Encoded by TCT"
FT Misc-difference 52
FT /note= "Encoded by TCC"
FT Misc-difference 53
FT /note= "Encoded by AGC"
FT Misc-difference 54
FT /note= "Encoded by AAC"
FT Misc-difference 56
FT /note= "Encoded by GTG"
FT Misc-difference 79
FT /note= "Encoded by CGG"
FT Misc-difference 81
FT /note= "Encoded by GAC"
FT Misc-difference 135
FT /note= "Encoded by CAG"
FT Misc-difference 136
FT /note= "Encoded by GCC"
FT Misc-difference 137
FT /note= "Encoded by GGC"
FT Misc-difference 138
FT /note= "Encoded by CAC"
FT Misc-difference 140
FT /note= "Encoded by GCC"
FT Misc-difference 141
FT /note= "Encoded by AGC"
FT Misc-difference 163
FT /note= "Encoded by CGG"
FT Misc-difference 164
FT /note= "Encoded by TCT"
FT Misc-difference 165
FT /note= "Encoded by GAC"
FT Misc-difference 166
FT /note= "Encoded by AAT"
FT Misc-difference 168

```

FT      /note= "Encoded by GTC"
FT      Misc-difference 169
FT      /note= "Encoded by CGG"
XX
XX      WO200152620-A2.
XX
XX      26-JUL-2001.
XX
XX      19-JAN-2001; 2001WO-US01817.
XX
XX      21-JAN-2000; 2000US-0177468.
XX      21-JUL-2000; 2000US-0620897.
XX
XX      (SCRI ) SCRIPPS RES INST.
XX      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX      Barbas CF, Stege JT, Guan X, Dalmia B;
XX
XX      WPI; 2001-465325/50.
XX      N-PSDB; AAD11588.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
XX      expression and metabolic pathways in plants, e.g. for treating in the
XX      plant cells a disorder that is associated with abnormal expression of
XX      the target gene -
XX
XX      Example 4; Fig 6; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the
XX      expression of a target gene in plant cells. The method involves
XX      providing plant cells with a zinc finger protein (ZFP) which is
XX      capable of specifically binding to a target nucleotide sequence
XX      or its complementary strand within a target gene and allowing the
XX      ZFP binding to the target nucleotide sequence, where the expression
XX      of the target gene in the plant cells is modulated. The ZFP and
XX      fusions of the ZFP proteins are useful for modulating or regulating
XX      gene expression and metabolic pathways in plants. The ZFP, fusion
XX      proteins and methods are useful in plant and agricultural technology.
XX      The method is useful particularly for treating a disorder in the plant
XX      cells, where the disorder is associated with abnormal expression of
XX      the target gene. The present sequence is Zinc finger protein (ZFP)
XX      m1 protein.
XX
XX      Query Match      73.4%; Score 45.5; DB 22; Length 184;
XX      Best Local Similarity 42.9%; Pred. No. 5;
XX      Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
XX
XX      1 CPEC-----XHQTH 10
XX      |||||
XX      154 CPEGKFSQAGHLASHQTH 174

RESULT 12
AAE06002
ID      AAE06002 standard; Protein: 184 AA.
XX
XX      AAE06002;
XX
XX      24-SEP-2001 (first entry)
XX
XX      Zinc finger protein (ZFP) m2 protein.
XX
XX      Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
XX      modulation; plant technology; agriculture.
XX
XX      Unidentified.
XX
XX      WO200152620-A2.
XX
XX      26-JUL-2001.
XX

```

```

PF      19-JAN-2001; 2001WO-US01817.
XX
XX      21-JAN-2000; 2000US-0177468.
XX      21-JUL-2000; 2000US-0620897.
XX
XX      (SCRI ) SCRIPPS RES INST.
XX      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX      Barbas CF, Stege JT, Guan X, Dalmia B;
XX
XX      WPI; 2001-465325/50.
XX      N-PSDB; AAD11589.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
XX      expression and metabolic pathways in plants, e.g. for treating in the
XX      plant cells a disorder that is associated with abnormal expression of
XX      the target gene -
XX
XX      Example 4; Fig 6; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the
XX      expression of a target gene in plant cells. The method involves
XX      providing plant cells with a zinc finger protein (ZFP) which is
XX      capable of specifically binding to a target nucleotide sequence
XX      or its complementary strand within a target gene and allowing the
XX      ZFP binding to the target nucleotide sequence, where the expression
XX      of the target gene in the plant cells is modulated. The ZFP and
XX      fusions of the ZFP proteins are useful for modulating or regulating
XX      gene expression and metabolic pathways in plants. The ZFP, fusion
XX      proteins and methods are useful in plant and agricultural technology.
XX      The method is useful particularly for treating a disorder in the plant
XX      cells, where the disorder is associated with abnormal expression of
XX      the target gene. The present sequence is zinc finger protein (ZFP)
XX      m2 protein.
XX
XX      Query Match      73.4%; Score 45.5; DB 22; Length 184;
XX      Best Local Similarity 42.9%; Pred. No. 5;
XX      Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
XX
XX      1 CPEC-----XHQTH 10
XX      |||||
XX      126 CPEGKFSQAGHLASHQTH 146

RESULT 13
AAB47899
ID      AAB47899 standard; Protein: 186 AA.
XX
XX      AAB47899;
XX
XX      16-MAY-2002 (first entry)
XX
XX      Zinc finger DNA binding domain, B3B.
XX
XX      Polymerase chain reaction; primer; amplify; PCR; reporter;
XX      polypeptide gene switch; ligand binding domain; LBD; transcription;
XX      nuclear hormone receptor; gene therapy; regulation; expression.
XX
XX      Synthetic.
XX
XX      Key      Location/Qualifiers
XX      Misc-difference 1
XX      /note= "Unspecified amino acid encoded by NNG"
XX      Misc-difference 186
XX      /note= "Unspecified amino acid encoded by CNN"
XX
XX      WO200206463-A2.
XX
XX      24-JAN-2002.
XX
XX      16-JUL-2001; 2001WO-EP08190.
XX

```

XX PR 18-JUL-2000; 2000US-0619063.
 XX PA (NOVS) NOVARTIS AG.
 XX PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 XX PA (SCRI) SCRIPPS RES INST.
 XX PI Beerli R, Schopfer U, Barbas CF;
 XX WPI; 2002-179793/23.
 XX N-PSDB; AAI72474.
 XX Single chain, monomeric polypeptide gene switches for regulating gene
 function, comprise two ligand binding domains derived from nuclear
 hormone receptors operatively linked to one functional domain -
 XX Disclosure: Fig 4: 63pp; English.
 XX This sequence represents the zinc finger DNA binding domain B3b. This
 protein could be used in the construction of a non-naturally occurring
 polypeptide gene switch comprising two ligand binding domains (LBDs)
 derived from nuclear hormone receptors operatively linked to a first
 functional domain. The gene switch is useful for regulating the function
 of a target nucleotide that contains a defined sequence, by exposing the
 target nucleotide to the polypeptide gene switch in the presence of a
 ligand that binds one of the LBDs of the polypeptide, where the DNA
 binding domain of the polypeptide binds the defined sequence, or the
 functional domain of the polypeptide alters the function of the target
 nucleotide. The gene switch is also useful in the field of gene therapy
 and as a regulator of gene expression or transcription. The advantage of
 the gene switches of the invention over existing gene switches is the
 need for only a single molecular switch and a single expression vector
 for production of that switch.
 XX Sequence 186 AA;
 Query Match 73.4%; Score 45.5; DB 23; Length 186;
 Best Local Similarity 42.9%; Pred. No. 5.1;
 Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
 QY 1 CPEC-----XHQRT 10
 DB 127 CPECCKSFQAGHLASHQRT 147
 RESULT 14
 ABU66292
 ID ABU66292 standard; Protein; 186 AA.
 XX AC ABU66292;
 XX DT 23-MAY-2003 (first entry)
 XX DE Zinc finger binding domain B3b.
 XX Zinc finger binding domain; gene switch; gene regulation.
 XX Unidentified.
 OS
 XX US2002168714-A1.
 XX PD 14-NOV-2002.
 XX 18-JUL-2001; 2001US-0908153.
 XX 18-JUL-2000; 2000US-0325747.
 XX (SCRI) SCRIPPS RES INST.
 XX Barbas CF, Beerli R, Schopfer U;
 XX WPI; 2003-328405/31.
 XX N-PSDB; ACA02218.

XX PT Novel polypeptide gene switch useful for regulating gene function,
 comprises two ligand binding domains derived from nuclear hormone
 receptors operatively linked to a functional domain -
 XX Disclosure: Fig 4: 33pp; English.
 XX The invention relates to a non-naturally occurring polypeptide (or
 polypeptide gene switch) comprising two ligand binding domains derived
 from nuclear hormone receptors operatively linked to a first functional
 domain. The polypeptide is useful for regulating the function of a target
 nucleotide that contains a defined sequence, by exposing the target
 nucleotide to the polypeptide in the presence of a ligand that binds one
 of the ligand binding domains of the polypeptide, where the DNA binding
 domain of the polypeptide binds the defined sequence or alters the
 function of the target nucleotide. The gene switches can be produced
 using a single molecular switch and a single expression vector. The
 present sequence represents the amino acid sequence of a zinc finger
 binding domain.
 XX Sequence 186 AA;
 Query Match 73.4%; Score 45.5; DB 24; Length 186;
 Best Local Similarity 42.9%; Pred. No. 5.1;
 Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
 QY 1 CPEC-----XHQRT 10
 DB 127 CPECCKSFQAGHLASHQRT 147
 RESULT 15
 ABP51407
 ID ABP51407 standard; Protein; 200 AA.
 XX AC ABP51407;
 XX DT 03-SEP-2002 (first entry)
 XX DE Human MDDT SEQ ID NO 429.
 XX Human; MDDT; disease detection and treatment molecule polynucleotide;
 KW proliferative disorder; hepatitis; psoriasis; cancer; AIDS;
 KW autoimmune disorder; inflammatory disorder; allergy; multiple sclerosis;
 KW rheumatoid arthritis; transgenic; gene therapy; antiarteriosclerotic;
 KW hepatotropic; antiinflammatory; antipsoriatic; cytostatic; anti-HIV;
 KW antiallergic; antianaemic; antiasthmatic; antiatherosclerotic; antigout;
 KW neuroprotective; antirheumatic; antiarthritic.
 XX OS Homo sapiens.
 XX WO200240715-A2.
 XX PD 23-MAY-2002.
 XX 06-SEP-2001; 2001WO-US27628.
 XX 06-SEP-2000; 2000US-230505P.
 XX 06-SEP-2000; 2000US-230514P.
 XX 06-SEP-2000; 2000US-230515P.
 XX 06-SEP-2000; 2000US-230517P.
 XX 06-SEP-2000; 2000US-230518P.
 XX 06-SEP-2000; 2000US-230519P.
 XX 06-SEP-2000; 2000US-230595P.
 XX 06-SEP-2000; 2000US-230597P.
 XX 06-SEP-2000; 2000US-230598P.
 XX 06-SEP-2000; 2000US-230599P.
 XX 06-SEP-2000; 2000US-230610P.
 XX 06-SEP-2000; 2000US-230865P.
 XX 06-SEP-2000; 2000US-230988P.
 XX 06-SEP-2000; 2000US-230989P.
 XX 07-SEP-2000; 2000US-230951P.
 XX 07-SEP-2000; 2000US-231163P.

```

PR 07-SEP-2000; 2000US-231167P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Jackson S, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Momiya MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
XX Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX WPI; 2002-527544/56.
DR N-PSDB; ABQ72624.
XX
XX Novel human disease detection and treatment polypeptide, useful in
PT diagnosis, prevention or treatment of cell proliferative disorders e.g.
PT arteriosclerosis, cirrhosis and an autoimmune/inflammatory disorder
PT e.g. AIDS
XX
XX Claim 14; Page 565; 618pp; English.
XX
XX The invention relates to an isolated human disease detection and
CC treatment (MDPT) polypeptide (I) selected from a polypeptide having a
CC sequence selected from 254 sequences (ABP51231-ABP51484) given in the
CC specification, a naturally occurring polypeptide comprising a sequence
CC having at least 90% identity to (I) or a biologically active or
CC immunogenic fragment of (I). (I) is useful for screening a compound for
CC effectiveness as an agonist or antagonist, for screening a compound that
CC specifically binds (I) or modulates the activity of (I), and for
CC preparing a polyclonal or monoclonal antibody by hybridoma technology.
CC Nucleic acids (II) (ABQ72449-ABQ72700) encoding (I) are useful for
CC screening a compound for effectiveness in altering expression of a target
CC polynucleotide comprising. Oligonucleotides and antibodies are useful for
CC detecting MDPT in a sample or for assessing toxicity of a test compound,
CC in a diagnostic test for a condition or a disease associated with the
CC expression of MDPT in a biological sample, for detecting (I) in a sample,
CC and for purifying (I) from a sample. A composition comprising (I), an
CC agonist or antagonist is useful for treating a disease or condition
CC associated with decreased or increased expression of functional MDPT.
CC (I) or (II) are useful for diagnosing, treating or preventing disorders
CC associated with aberrant expression of MDPT, where the disorders are
CC selected from a cell proliferative disorder such as arteriosclerosis,
CC cirrhosis, hepatitis, psoriasis, and cancer and an
CC autoimmune/inflammatory disorder such as AIDS, Addison's disease,
CC allergy, anaemia, asthma, atherosclerosis, gout, multiple sclerosis or
CC rheumatoid arthritis. (II) are useful for creating knockin humanised
CC animals or transgenic animals to model human diseases, in somatic or
CC germ-line gene therapy, to generate a transcript image of a tissue or cell
CC type, for detecting differences in the chromosomal location due to
CC translocation or inversion among normal, carrier or affected individuals
CC and as hybridisation probes for mapping naturally occurring genomic
CC sequences.
XX
XX Sequence 200 AA;
XX
Query Match 73.48; Score 45.5; DB 23; Length 200;
Best Local Similarity 42.9%; Pred. No. 5.5;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
QY 1 CPEC-----XHQRTH 10
DB |||||
27 CPECGRSFSKALKTHQRTH 47

```

Search completed: October 1, 2003, 10:41:47
 Job time : 64.8889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:34:26 ; Search time 21.9444 Seconds
(without alignments)
43.824 Million cell updates/sec

Title: US-10-057-890A-3
Perfect score: 62
Sequence: 1 CPECXHQTH 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.5	73.4	614	2 JH0500	zinc finger protei
2	44.5	71.8	196	2 S06555	finger protein (cl
3	44.5	71.8	462	2 S10397	finger protein kox
4	44.5	71.8	485	2 A0751	finger protein MZF
5	44.5	71.8	540	2 T14748	hypothetical prote
6	44.5	71.8	1350	2 S00847	finger protein - A
7	41.5	66.9	109	2 H45193	zinc finger protei
8	41	66.1	191	2 A39492	rubrerythrin - Des
9	41	66.1	746	2 S50226	CSD3 protein - yea
10	41	66.1	1306	2 T13592	hypothetical prote
11	40.5	65.3	82	2 E29942	developmental cont
12	40.5	65.3	224	2 S06545	finger protein (cl
13	40.5	65.3	509	2 JH0501	phnA protein NMB00
14	40	64.5	109	2 G81244	phnA protein homol
15	40	64.5	109	2 F82023	conserved hypotet
16	40	64.5	111	2 A11021	conserved hypotet
17	40	64.5	112	2 E95186	phnA protein limpo
18	40	64.5	112	2 F98053	hypothetical prote
19	40	64.5	113	2 A83631	conserved hypotet
20	39.5	63.7	424	2 S15917	developmental cont
21	39	62.9	111	2 AC1121	conserved hypotet
22	39	62.9	111	2 AF1481	conserved hypotet
23	39	62.9	113	2 AB0333	conserved hypotet
24	39	62.9	190	2 AB2370	DNA topoisomerase
25	39	62.9	232	2 H70029	hypothetical prote
26	39	62.9	273	2 T51145	nucleic acid bindi
27	39	62.9	381	2 B97182	acyl-protein synth
28	38.5	62.1	112	2 S06562	finger protein (cl
29	38.5	62.1	390	2 S06222	finger protein sna

30	38.5	62.1	459	2 T37704	zinc-finger protei
31	38.5	62.1	459	2 T43538	zinc finger protei
32	38.5	62.1	634	2 T47156	hypothetical prote
33	38	61.3	76	2 AH1584	B. subtilis Y1ai p
34	38	61.3	80	2 AE1208	B. subtilis Y1ai p
35	38	61.3	165	2 B95891	conserved hypotet
36	38	61.3	191	2 AE2559	hypothetical prote
37	38	61.3	204	2 F64391	rubrerythrin - Met
38	38	61.3	216	2 AH2547	hypothetical prote
39	38	61.3	343	2 S09272	Ig alpha chain C r
40	38	61.3	392	2 PC1144	finger protein Znc
41	38	61.3	412	2 D30341	G protein-coupled
42	38	61.3	420	2 S36444	hygromycin phospho
43	38	61.3	674	2 S52240	DNA ligase (NAD) (
44	38	61.3	676	2 A40363	DNA ligase (NAD) (
45	38	61.3	698	2 S25409	transcription fact

ALIGNMENTS

RESULT 1

JH0500
zinc finger protein 29 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Dec-2000
C:Accession: JH0500
R:Denny, P.; Ashworth, A.
Gene 106, 221-227, 1991
A:Title: A zinc finger protein-encoding gene expressed in the post-meiotic phase of
A:Reference number: JH0500; MUID:92039080; PMID:1937051
A:Accession: JH0500
A:Molecule type: mRNA
A:Residues: 1-614 <DEN>
A:Cross-references: GB:X55126; NID:g55470; PIDN:CAA38920.1; PID:g55471
A:Experimental source: Testis
C:Keywords: DNA binding; zinc finger
F:217-244/Region: zinc finger
F:245-272/Region: zinc finger
F:273-300/Region: zinc finger
F:301-328/Region: zinc finger
F:329-356/Region: zinc finger
F:357-384/Region: zinc finger
F:385-412/Region: zinc finger
F:413-440/Region: zinc finger
F:441-468/Region: zinc finger
F:469-496/Region: zinc finger
F:497-524/Region: zinc finger
F:525-552/Region: zinc finger
F:553-580/Region: zinc finger
F:581-608/Region: zinc finger

Query Match 73.4%; Score 45.5; DB 2; Length 614;

Best Local Similarity 42.9%; Pred. No. 8.8;

Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10

|||||

Db 588 CPECCKGFSNSSNFTHQTH 608

RESULT 2

S06555

finger protein (clone X1COF26) - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 16-Jul-1999

R:Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Posting, A.; Ki

J. Mol. Biol. 208, 639-659, 1989

A:Title: Second-order repeats in Xenopus laevis finger proteins.

A:Reference number: S05632; MUID:90040698; PMID:2509712

A:Accession: S06555

A>Status: not compared with conceptual translation

```
A:Molecule type: mRNA
A:Residues: 1-196 <NIE>
C:Superfamily: LIM metal-binding repeat homology
C:Keywords: DNA binding; zinc finger
F:39-95/Domain: LIM metal-binding repeat homology <LIM>

Query Match 71.8%; Score 44.5; DB 2; Length 196;
Best Local Similarity 42.9%; Pred. No. 5.5;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQRT 10
    ||||
Db 120 CPECCKKFSVKNLKHQRT 140
    |||||

RESULT 3
S10397
finger protein kox1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 01-Dec-2000
C:Accession: S10397
submitted to the EMBL Data Library, March 1990
A:Reference number: S10397
A:Molecule type: mRNA
A:Residues: 1-462 <THI>
A:Cross-references: EMBL:X52332
C:Genetics:
A:Gene: GDB:2NF10; KOX1
A:Cross-references: GDB:125317; OMIM:194538
A:Map position: 12q24.33-12q24.33
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C:Keywords: DNA binding; zinc finger

Query Match 71.8%; Score 44.5; DB 2; Length 462;
Best Local Similarity 42.9%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQRT 10
    ||||
Db 268 CPECCKKFSVKNLKHQRT 288
    |||||

RESULT 4
A40751
finger protein MZF1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 01-Dec-2000
C:Accession: A40751
F:Thomas, R.; Collins, S.J.; Hickstein, D.; Raskind, W.; Deaven, L.L.; O'Hara, P.; Hagen
J. Biol. Chem. 266, 14183-14187, 1991
A:Title: A retinoic acid-responsive human zinc finger gene, MZF-1, preferentially expres
A:Reference number: A40751; MUID:91317761; PMID:1860835
A:Accession: A40751
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-485 <HRO>
A:Cross-references: GB:M58297; NID:g189043; PID:g189044
C:Genetics:
A:Gene: GDB:2NF42; MZF-1
A:Cross-references: GDB:125898; OMIM:194550
A:Map position: 19q13.2-19q13.4
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 71.8%; Score 44.5; DB 2; Length 485;
Best Local Similarity 42.9%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQRT 10
    ||||
Db 350 CPECCKKFSVKNLKHQRT 370
    |||||

A:Molecule type: mRNA
A:Residues: 1-196 <NIE>
C:Superfamily: LIM metal-binding repeat homology
C:Keywords: DNA binding; zinc finger
F:39-95/Domain: LIM metal-binding repeat homology <LIM>

Query Match 71.8%; Score 44.5; DB 2; Length 196;
Best Local Similarity 42.9%; Pred. No. 5.5;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQRT 10
    ||||
Db 120 CPECCKKFSVKNLKHQRT 140
    |||||

RESULT 5
T14748
hypothetical protein DKF2p434G1621.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C:Accession: T14748
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: Z18179
A:Accession: T14748
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-540 <BLU>
A:Cross-references: EMBL:AL110188
A:Experimental source: adult testis; clone DKF2p434G1621
C:Genetics:
A:Note: DKF2p434G1621.1
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 71.8%; Score 44.5; DB 2; Length 540;
Best Local Similarity 42.9%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQRT 10
    ||||
Db 346 CPECCKKFSVKNLKHQRT 366
    |||||

RESULT 6
S00647
finger protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 31-Dec-1993
C:Accession: S00647
R;Ruiz i Altaba, A.; Perry-O'Keefe, H.; Melton, D.A.
EMBO J. 6, 3065-3070, 1987
A:Title: Xfin: an embryonic gene encoding a multifingered protein in Xenopus.
A:Reference number: S00647; MUID:88082679; PMID:2826129
A:Accession: S00647
A:Molecule type: mRNA
A:Residues: 1-1350 <RUI>
A:Cross-references: EMBL:X06021
A:Note: it is uncertain whether Met-1, Met-11, Met-16, Met-38, or Met-39 is the init
C:Genetics:
A:Gene: fin
C:Keywords: DNA binding; zinc finger

Query Match 71.8%; Score 44.5; DB 2; Length 1350;
Best Local Similarity 42.9%; Pred. No. 22;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQRT 10
    ||||
Db 808 CPECCKKFSVKNLKHQRT 828
    |||||

RESULT 7
H45193
zinc finger protein ZNF65 (C2-H2 zinc finger motif) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Jan-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: H45193
R;Lichter, P.; Bray, P.; Ried, T.; Dawid, I.B.; Ward, D.C.
Genomics 13, 999-1007, 1992
A:Title: Clustering of C2-H2 zinc finger motif sequences within telomeric and fragile
A:Reference number: A43284; MUID:92372070; PMID:1505991
A:Accession: H45193
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-109 <LIC>
A:Note: sequence extracted from NCBI backbone (NCBIN:111681, NCBIP:111682)
```

Query Match 66.9%; Score 41.5; DB 2; Length 109;
Best Local Similarity 38.1%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPECXHQTH 10
|||
DB 63 CPDCGKCFSSNLVQHQRTH 83

RESULT 8
A39492
rubrerythrin - Desulfovibrio vulgaris
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 29-Sep-1999
C:Species: Desulfovibrio vulgaris
A:Accession: A41191; A39492; S29385
R:Prickril, B.C.; Kurtz J.F.; D.M.; LeGall, J.
Biochemistry 30, 11118-11123, 1991
A:Title: Cloning and sequencing of the gene for rubrerythrin from Desulfovibrio vulgaris
A:Reference number: A41191; MUID:92031575; PMID:1932032
A:Molecule type: DNA
A:Accession: A41191
A:Residues: 1-191 <PR>
A:Cross-references: GB:U82323; GB:M77011; NID:gl769569; PIDN:AAB39991.1; PID:gl769571
R:Van Beem, J.J.; Van Driessche, G.; Liu, M.Y.; LeGall, J.
J. Biol. Chem. 266, 20645-20653, 1991
A:Title: The primary structure of rubrerythrin, a protein with inorganic pyrophosphatase
A:Reference number: A39492; MUID:92041914; PMID:1657933
A:Accession: A39492
A:Molecule type: protein
A:Residues: 1-191 <VAN>
R:Pierik, A.J.; Wolbert, R.B.G.; Portier, G.L.; Verhagen, M.F.J.M.; Hagen, W.R.
Eur. J. Biochem. 212, 237-245, 1993
A:Title: Nigerythrin and rubrerythrin from Desulfovibrio vulgaris each contain two mononucleotide binding sites
A:Reference number: S29385; MUID:93185629; PMID:8383040
A:Accession: S29385
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-15 <PIE>
C:Complex: homodimer
C:Superfamily: rubrerythrin; rubredoxin homology
C:Keywords: homodimer; iron; metalloprotein
F:158-184/Domain: rubredoxin homology KRUB>
F:158,161,174,177/Binding site: iron (Cys) #status predicted

Query Match 66.1%; Score 41; DB 2; Length 191;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CPECXHQTH 10
|||
DB 174 CPACAPKRAH 183

RESULT 9
S50226
CSD3 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J0838; protein YJL099w
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 29-Oct-1999
C:Accession: S50226; S53388; S57369; S56877
R:Osmond, B.C.
submitted to the EMBL Data Library, October 1994
A:Description: Sequence analysis and disruption of the CSD3 Gene of Saccharomyces cerevisiae
A:Reference number: S50226
A:Molecule type: DNA
A:Accession: S50226
A:Residues: 1-746 <OSH>
A:Cross-references: EMBL:U15603; NID:g559796; PIDN:AAA50840.1; PID:g559797
R:Rasmussen, S.W.
submitted to the EMBL Data Library, February 1995
A:Description: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and
A:Reference number: S53376

A:Accession: S53388
A:Molecule type: DNA
A:Residues: 1-746 <RAS>
A:Cross-references: EMBL:X85021; NID:g728698; PIDN:CAA59395.1; PID:g728711
R:Rasmussen, S.W.
Yeast 11, 873-883, 1995
A:Title: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and CS
A:Reference number: S57357; MUID:96090136; PMID:7483851
A:Accession: S57369
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-746 <RAF>
A:Cross-references: EMBL:X85021; NID:g728698; PIDN:CAA59395.1; PID:g728711
R:Rasmussen, S.W.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56876
A:Accession: S56877
A:Molecule type: DNA
A:Residues: 1-746 <RAW>
A:Cross-references: EMBL:Z49374; NID:gl008277; PIDN:CAA89394.1; PID:gl008278; MIPS:Y
C:Genetics:
A:Gene: SGD:C56; CSD3
A:Cross-references: MIPS:YJL099w; SGD:S0003635
A:Map position: 10L

Query Match 66.1%; Score 41; DB 2; Length 746;
Best Local Similarity 44.4%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

QY 1 CP-----ECXHQRTH 10
|||
DB 43 CPTMRQFGECLEHNRTH 60

RESULT 10
T13592
hypothetical protein 66A1.2 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13592
R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demalle, J.G.
submitted to the EMBL Data Library, October 1998
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: z17692
A:Accession: T13592
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1306 <FER>
A:Cross-references: EMBL:AL031227; NID:ei330103; PID:ei316854; PIDN:CAA20225.1
C:Genetics:
A:Cross-references: FlyBase:FBgn0025388
A:Introns: 676/1; 1011/3; 1254/1
A>Note: EG:66A1.2

Query Match 66.1%; Score 41; DB 2; Length 1306;
Best Local Similarity 66.7%; Pred. No. 74;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CPECXHQTH 9
|||
DB 995 CAQCGHQRT 1003

RESULT 11
E29942
developmental control protein Krox-8 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 05-Nov-1999
C:Accession: E29942
R:Chavrier, P.; Lemaire, P.; Revelant, O.; Bravo, R.; Charnay, P.
Mol. Cell. Biol. 8, 1319-1326, 1988
A:Title: Characterization of a mouse multigene family that encodes zinc finger struc

A:Reference number: A29942; MUID:98216608; PMID:2452975
A:Accession: E29942
A:Molecule type: DNA
A:Residues: 1-82 <CHA>
A:Cross-references: GB:M19440; NID:q198616; PIDN:AAA39388.1; PID:g554181
C:Keywords: DNA binding; zinc finger

Query Match 65.3%; Score 40.5; DB 2; Length 82;
Best Local Similarity 38.1%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 2; Indels 11; Gaps 1;

QY 1 CPEC-----XHORTH 10
||||| |
Db 9 CPEGCKFSQRSNLIAHRTH 29
||||| |

RESULT 12
S06545
finger protein (clone XLOC6.1) - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 31-Dec-1993
C:Accession: S06545
R:Nietfeld, W.; El-Baradi, T.; Mentzel, H.; Pieler, T.; Koester, M.; Poeting, A.; Knoechel, J. Mol. Biol. 208, 639-659, 1989
A:Title: Second-order repeats in Xenopus laevis finger proteins.
A:Reference number: S05632; MUID:90040698; PMID:2505712
A:Accession: S06545
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-224 <NIE>
C:Keywords: DNA binding; zinc finger

Query Match 65.3%; Score 40.5; DB 2; Length 224;
Best Local Similarity 38.1%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHORTH 10
||||| |
Db 148 CPEGCKFTQSMQLIRHRTH 168
||||| |

RESULT 13
RH0501
zinc finger protein Ckr1 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Dec-2000
C:Accession: JH0501; S19007
R:Benn, A.; Antoine, M.; Beug, H.; Niessing, J.
Gene 106, 207-212, 1991
A:Title: Primary structure and expression of a chicken cDNA encoding a protein with zinc finger
A:Reference number: JH0501; MUID:92039078; PMID:1840543
A:Accession: JH0501
A:Molecule type: mRNA
A:Residues: 1-509 <BEN>
A:Cross-references: GB:X15538; NID:g63228; PIDN:CAA33545.1; PID:g63229
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology
C:Keywords: DNA binding; zinc finger
F:113-140/Region: zinc finger
F:141-168/Region: zinc finger
F:169-196/Region: zinc finger
F:197-224/Region: zinc finger
F:225-252/Region: zinc finger
F:303-330/Region: zinc finger
F:331-358/Region: zinc finger
F:359-386/Region: zinc finger
F:387-414/Region: zinc finger
F:415-442/Region: zinc finger

Query Match 65.3%; Score 40.5; DB 2; Length 509;
Best Local Similarity 38.1%; Pred. No. 44;
Matches 8; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHORTH 10
||||| |

Db 143 CPEGESYSQSHLVHRTH 163
||||| |

RESULT 14
G81244
pHnA protein NMB0037 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: G81244
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: G81244
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-109 <TET>
A:Cross-references: GB:AE002364; GB:AE002098; NID:g7225269; PIDN:AAF40508.1; PID:g722
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0037

Query Match 64.5%; Score 40; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CPECXHQ 7
||||| |
Db 23 CPECAHE 29
||||| |

RESULT 15
F82023
PhnA protein homolog NMA0283 [imported] - Neisseria meningitidis (strain 22491 serogr
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: F82023
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mc ; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: F82023
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-109 <PAR>
A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83590.1; PID:g737
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: phnA; NMA0283

Query Match 64.5%; Score 40; DB 2; Length 109;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CPECXHQ 7
||||| |
Db 23 CPECAHE 29
||||| |

Search completed: October 1, 2003, 10:47:46
Job time : 23.9444 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:06:46 ; Search time 11.6667 Seconds
(without alignments)
40.309 Million cell updates/sec

Title: US-10-057-890A-3
Perfect score: 62
Sequence: 1 CPEXHQTH 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues 127863
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45.5	73.4	614	1 ZF29_MOUSE	Q07230 mus musculus
2	44.5	71.8	196	1 Z026_XENLA	P18746 xenopus lae
3	44.5	71.8	439	1 ZN10_HUMAN	P21506 homo sapien
4	44.5	71.8	489	1 ZN71_HUMAN	Q9nqz8 homo sapien
5	44.5	71.8	512	1 Z169_HUMAN	Q14929 homo sapien
6	44.5	71.8	520	1 ZF94_MOUSE	Q9z1d9 mus musculus
7	44.5	71.8	536	1 ZF94_MOUSE	Q9z2k3 rattus norv
8	44.5	71.8	734	1 ZN42_HUMAN	P28698 homo sapien
9	44.5	71.8	1350	1 XFIN_XENLA	P08045 xenopus lae
10	41.5	66.9	481	1 Z358_HUMAN	Q9nq07 homo sapien
11	41.5	66.9	548	1 Z382_HUMAN	Q968r6 homo sapien
12	41.5	66.9	683	1 Z263_HUMAN	O14978 homo sapien
13	41	66.1	191	1 RUBY_DESVH	P24931 desulfovibr
14	41	66.1	746	1 CHS6_YEAST	P40955 saccharomyc
15	40.5	65.3	82	1 ZF13_MOUSE	P10754 mus musculus
16	40.5	65.3	224	1 Z061_XENLA	P18750 xenopus lae
17	40.5	65.3	504	1 Z205_HUMAN	Q95201 homo sapien
18	40.5	65.3	509	1 ZKRI_CHICK	P30373 gallus gall
19	40.5	65.3	606	1 Z214_HUMAN	Q9u159 homo sapien
20	39.5	63.7	424	1 ZFPI_MOUSE	P08042 mus musculus
21	39	62.9	108	1 RPD3_HUMAN	Q9y2y1 homo sapien
22	39	62.9	689	1 PRD3_HUMAN	Q9nq07 homo sapien
23	39	62.9	1267	1 Y211_HUMAN	Q92610 homo sapien
24	38.5	62.1	112	1 ZG9_XENLA	P18738 xenopus lae
25	38.5	62.1	216	1 Z434_HUMAN	Q9nq65 homo sapien
26	38.5	62.1	390	1 SNA1_DROME	P08044 drosophila
27	38.5	62.1	459	1 ZPRL_SCHPO	O13724 schizosacch
28	38.5	62.1	648	1 Z202_HUMAN	Q95125 homo sapien
29	38.5	62.1	810	1 Z33A_HUMAN	Q06730 homo sapien
30	38	61.3	204	1 RUBY_METJA	O58144 methanococc
31	38	61.3	412	1 AA2A_CANFA	P11617 canis faml
32	38	61.3	670	1 DNLI_THEFI	Q9zh10 thermus fil
33	38	61.3	674	1 DNLI_THESC	P49422 thermus sco

34	38	61.3	676	1 DNLI_THETH	P26996 thermus the
35	38	61.3	771	1 ZN06_HUMAN	Q9y462 homo sapien
36	37.5	60.5	56	1 ZN21_HUMAN	P17025 homo sapien
37	37.5	60.5	168	1 ZG7_XENLA	P18735 xenopus lae
38	37.5	60.5	196	1 Z064_XENLA	P18732 xenopus lae
39	37.5	60.5	393	1 ZF46_MOUSE	Q03309 mus musculu
40	37.5	60.5	394	1 Z200_HUMAN	P98182 homo sapien
41	37.5	60.5	692	1 Z180_HUMAN	Q9u1w8 homo sapien
42	37.5	60.5	697	1 HKRI_HUMAN	P10072 homo sapien
43	37.5	60.5	738	1 ZN84_HUMAN	P51523 homo sapien
44	37.5	60.5	751	1 Z337_HUMAN	Q9y3m9 homo sapien
45	37.5	60.5	792	1 Z328_HUMAN	Q8wxb4 homo sapien

ALIGNMENTS

RESULT 1					
ZF29_MOUSE					
ID	ZF29_MOUSE	STANDARD;	PRT;	614 AA.	
AC	Q07230;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Zinc finger protein 29 (Zfp-29).				
GN	ZFP29 OR ZFP-29.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID:10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Testis;				
RX	MEDLINE-92039080; PubMed-1937051;				
RA	Denny P., Ashworth A.;				
RT	"A zinc finger protein-encoding gene expressed in the post-meiotic				
RT	Phase of spermatogenesis.";				
RL	Gene 106:221-227(1991).				
CC	!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION DURING				
CC	THE POST-MEIOTIC STAGES OF SPERMATOGENESIS.				
CC	!- SUBCELLULAR LOCATION: Nuclear.				
CC	!- TISSUE SPECIFICITY: IN THE ADULT, PREDOMINANTLY FOUND IN				
CC	SPERMATIDS. ALSO PRESENT IN THE EMBRYO.				
CC	!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-				
CC	FINGER PROTEINS.				
CC	!- SIMILARITY: Contains 1 SCAN box domain.				
CC	!- SIMILARITY: Contains 14 C2H2-type zinc fingers.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; X55126; CAA38920.1; -				
DR	PIR; JH0500; JH0500.				
DR	HSSP; P08047; 1SP2.				
DR	MGD; MGI:99176; Zfp29.				
DR	InterPro; IPR003309; Treg_SCAN.				
DR	InterPro; IPR007087; Znf_C2H2.				
DR	InterPro; IPR007086; Znf_C2H2_sub.				
DR	Pfam; PF02023; SCAN; 1.				
DR	Pfam; PF00096; zf-C2H2; 14.				
DR	PRINTS; PR00048; ZINCFINGER.				
DR	ProDom; PD000003; Znf_C2H2; 13.				
DR	SMART; SM00431; LER; 1.				
DR	SMART; SM00355; Znf_C2H2; 14.				
DR	PROSITE; PS00804; SCAN_BOX; 1.				
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14.				
DR	PROSITE; PS0157; ZINC_FINGER_C2H2_2; 14.				
DR	Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat;				
KW					

KW Developmental protein; Transcription regulation; Spermatogenesis.

FT DOMAIN 69 127
 FT ZN_FING 222 244
 FT ZN_FING 250 272
 FT ZN_FING 278 300
 FT ZN_FING 306 328
 FT ZN_FING 334 356
 FT ZN_FING 362 384
 FT ZN_FING 390 412
 FT ZN_FING 418 440
 FT ZN_FING 446 468
 FT ZN_FING 474 496
 FT ZN_FING 502 524
 FT ZN_FING 530 552
 FT ZN_FING 558 580
 FT ZN_FING 586 608
 SQ SEQUENCE 614 AA; 68714 MW; 738156A0A9413DF7 CRC64;

Query Match 73.4%; Score 45.5; DB 1; Length 614;
 Best Local Similarity 42.9%; Pred. No. 0.9;
 Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

OY 1 CPEC-----XHQTH 10
 DB 588 CPECCKGFSNSSNFTHQTH 608

RESULT 2

ID Z026_XENLA STANDARD; PRT; 196 AA.
 AC P18746;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Oocyte zinc finger protein XLCOF26 (Fragment).
 OS Xenopus laevis (African clawed frog)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-90040698; PubMed-2509712;
 RA Nietfeld W., El-Baradi T., Menzel H., Pieler T., Koester M.,
 RA Poetling A., Knoechel W.;
 RT "Second-order repeats in Xenopus laevis finger proteins.";
 RL J. Mol. Biol. 208:639-659(1989).
 DR PIR; S06555; S06555.
 DR HSSP; P08047; 1SP1.

DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 7.
 DR SMART; SM00355; Znf_C2H2; 7.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 7.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.

FT NON_TER 1 1
 FT ZN_FING 6 28
 FT ZN_FING 34 56
 FT ZN_FING 62 84
 FT ZN_FING 90 112
 FT ZN_FING 118 140
 FT ZN_FING 146 168
 FT ZN_FING 174 196
 FT NON_TER 196 196
 SQ SEQUENCE 196 AA; 22659 MW; A4417A032257D595 CRC64;

Query Match 71.8%; Score 44.5; DB 1; Length 196;
 Best Local Similarity 42.9%; Pred. No. 0.46;
 Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

OY 1 CPEC-----XHQTH 10
 DB 120 CPECCKKFSVKNLSLRKHQTH 140

RESULT 3
 ZN10_HUMAN STANDARD; PRT; 439 AA.
 AC P21506;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 10 (zinc finger protein KOX1) (Fragment).
 GN ZNF10 OR KOX1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-91145339; PubMed-2288909;
 RA Thiesen H.-J.;
 RT "Multiple genes encoding zinc finger domains are expressed in human T cells.";
 RL New Biol. 2:363-374(1990).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.

CC -1- SIMILARITY: Contains 1 KRAB domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)
 CC -----

DR EMBL; X52332; CAA36558.1; -;
 DR HSSP; P08047; 1SP2.
 DR TRANSFAC; T02283; -;
 DR Genew; HGNC:12879; ZNF10.
 DR MIM; 194538; -;

DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR007087; Znf_C2H2.
 DR InterPro: IPR007086; Znf_C2H2_sub.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00096; zf-C2H2; 9.
 DR PRINTS; PR00048; ZINCFINGER.
 DR SMART; SM00349; KRAB; 1.

DR SMART; SM00355; Znf_C2H2; 9.
 DR PROSITE; PS00805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 9.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat.

FT DOMAIN 14 85
 FT ZN_FING 159 181
 FT ZN_FING 187 209
 FT ZN_FING 215 237
 FT ZN_FING 243 265
 FT ZN_FING 271 293
 FT ZN_FING 299 321
 FT ZN_FING 327 349
 FT ZN_FING 355 377
 FT ZN_FING 383 405
 FT ZN_FING 410 432
 FT NON_TER 439 439
 SQ SEQUENCE 439 AA; 51220 MW; D772906259D2B78C CRC64;

Query Match 71.8%; Score 44.5; DB 1; Length 439;
 Best Local Similarity 42.9%; Pred. No. 0.98;
 Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

OY 1 CPEC-----XHQTH 10

DB 245 CPEGKSFQSTHLLHORTH 265
|||||
|||||
RESULT 4
N71L HUMAN STANDARD: PRT; 489 AA.
AC O9NC028: O15919; O9UC09: O9UC03;
DT 28-FEB-2003 (Rel. 41, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Endothelial zinc finger protein induced by tumor necrosis factor alpha
DE (zinc finger protein 71) (ZNF47).
GN ZNF71 OR ZNF71.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Umbilical vein endothelial cells;
RA MEDLINE=21319880; PubMed=11426589;
RA Matakai C., Murakami T., Umetani M., Wada Y., Ishii M., Tsutsumi S.,
RA Aburatani H., Hamakubo T., Kodama T.;
RT "A novel zinc finger protein mRNA in human umbilical vein endothelial
RT cells is profoundly induced by tumor necrosis factor alpha";
RL J. Atheroscler. Thromb. 7:97-103(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Sakaladasis G.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Scott D., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Danganan L.,
RA Erler A., Christensen M., Georgescu A., Avila J., Attix C.,
RA Andreise T., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Krommiller B., Arellano A., Sanders C.,
RA Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;
RT "Sequence analysis of a 3.6 Mb region in 19q13.4 between D19S891 to
RT ZNF134";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA MEDLINE=22388257; PubMed=1247932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 151-235 FROM N.A.
RX MEDLINE=92347859; PubMed=1639391;
RA Aubry M., Marineau C., Zhang F.R., Zahed L., Figlewicz D.,
RA Delattre O., Thomas G., de Jong P.J., Julien J.-P., Rouleau G.A.;
RT "Cloning of six new genes with zinc finger motifs mapping to short and
RT long arms of human acrocentric chromosome 22 (p and q11.2).";
RL Genomics 13:641-648(1992).
RN [5]
RP SEQUENCE OF 431-489 FROM N.A.

RC TISSUE=Placenta;
RX MEDLINE=92372070; PubMed=1505991;
RA Clachter P., Bray P., Ried T., David I.B., Ward D.C.;
RT "Clustering of C2-H2 zinc finger motif sequences within telomeric and
RT fragile site regions of human chromosomes.";
RL Genomics 13:999-1007(1992).
CC -|- FUNCTION: May be involved in transcriptional regulation.
CC -|- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -|- TISSUE SPECIFICITY: Highly expressed in placenta, followed by
CC brain, testis, pancreas, heart, small intestine, muscle, uterus,
CC prostate and peripheral blood leukocytes. Not detected in liver,
CC lung, colon, stomach, salivary and thyroid gland.
CC -|- INDUCTION: By tumor necrosis factor alpha (TNFalpha).
CC -|- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -|- SIMILARITY: Contains 13 C2H2-type zinc fingers.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF269249; AAF88036.1; -;
DR EMBL; AC007228; AAD23608.1; ALT_INIT.
DR EMBL; BC014280; AAH14280.1; -;
DR EMBL; X60074; CAC16174.1; -;
DR EMBL; M88357; AAC61315.1; -;
DR PIR; A42825; A42825.
DR HSP; P07248; 1PAA.
DR Genew; HGNC:13141; ZNF71.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF000096; zf-C2H2; 13.
DR PRINTS; PR00048; ZINC_FINGER.
DR PRODOM; PD000003; Znf_C2H2; 11.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 13.
DR Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT ZN_FING 130 152 C2H2-TYPE.
FT ZN_FING 158 180 C2H2-TYPE.
FT ZN_FING 186 208 C2H2-TYPE.
FT ZN_FING 214 236 C2H2-TYPE.
FT ZN_FING 242 264 C2H2-TYPE.
FT ZN_FING 270 292 C2H2-TYPE.
FT ZN_FING 298 320 C2H2-TYPE.
FT ZN_FING 326 348 C2H2-TYPE.
FT ZN_FING 354 376 C2H2-TYPE.
FT ZN_FING 382 404 C2H2-TYPE.
FT ZN_FING 410 432 C2H2-TYPE.
FT ZN_FING 438 460 C2H2-TYPE.
FT ZN_FING 466 488 C2H2-TYPE.
SQ SEQUENCE 489 AA; 54497 MW; 6B1156D21FF9295C CRC64;

Query Match 71.8%; Score 44.5; DB 1; Length 489;
Best Local Similarity 42.9%; Pred No. 1.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEGKSFQSTHLLHORTH 10
|||||
Db 272 CPEGKSFQSTHLLHORTH 292
|||||
RESULT 5
ID 2169_HUMAN STANDARD: PRT; 512 AA.
AC Q14929;
DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 169 (Fragment).
GN ZNF169.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97225201; PubMed=9071574;
RA Chidambaram A., Gailani M., Gorrard B., Stewart C., Goldstein A.,
RA Chakrabarti I., Bale A.E., Dean M.;
RT "Characterization of a YAC contig containing the NBCCS locus and a
RT novel kruppel-type zinc finger sequence on chromosome segment
RT 9q22.3.";
RL Genes Chromosomes Cancer 18:212-218(1997).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN KIDNEY, IS WEAKLY
CC EXPRESSED IN HEART, LIVER, SPLEEN, AND SMALL INTESTINE, AND IS NOT
CC EXPRESSED IN ADULT BRAIN OR SPINAL CORD.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U28251; AAA70188.1;
CC HSSP; P25490; 1UBD.
CC Genew; HGNC:12957; ZNF169.
CC MIM; 603404;
CC GO; GO:0005634; C:nucleus; NAS.
CC GO; GO:0003700; E:transcription factor activity; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC InterPro; IPR007087; Znf_C2H2.
CC Pfam; PF00096; zf-C2H2; 11.
CC SMART; SM00355; Znf_C2H2; 10.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 12.
CC Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT NON_TER 1
FT ZN_FING 144 166 C2H2-TYPE.
FT ZN_FING 172 194 C2H2-TYPE.
FT ZN_FING 200 222 C2H2-TYPE.
FT ZN_FING 228 250 C2H2-TYPE.
FT ZN_FING 311 333 C2H2-TYPE.
FT ZN_FING 339 361 C2H2-TYPE.
FT ZN_FING 367 389 C2H2-TYPE.
FT ZN_FING 395 417 C2H2-TYPE.
FT ZN_FING 449 473 C2H2-TYPE.
FT ZN_FING 479 504 C2H2-TYPE.
SQ SEQUENCE 512 AA; 57651 MW; 5B78DF763C63E9A2 CRC64;
Query Match 71.8%; Score 44.5; DB 1; Length 512;
Best Local Similarity 42.9%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
QY 1 CPEC-----XHQTH 10
Db 146 CPEGCGFCQSRDLKHQTH 166
RESULT 6
ZF94_MOUSE
ID ZF94_MOUSE PRT; 520 AA.
AC Q92ID9;

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 94 (Zfp-94).
GN ZFP94.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX Weisig H., Millan J.L.;
RA "A family of zinc finger genes is differentially expressed during
RA spermatogenesis".
RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 1 SCAN box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U62906; AAD00102.1;
CC HSSP; P08045; 1ZNF.
CC MGD; MGI:107610; Zfp94.
CC InterPro; IPR001909; KRAB.
CC InterPro; IPR003309; Treg_SCAN.
CC InterPro; IPR007087; Znf_C2H2.
CC InterPro; IPR007086; Znf_C2H2_sub.
CC Pfam; PF01352; KRAB; 1.
CC Pfam; PF02023; SCAN; 1.
CC Pfam; PF00096; zf-C2H2; 7.
CC PRINTS; PR00048; ZINCFINGER.
CC ProDom; PD000003; Znf_C2H2; 6.
CC SMART; SM00349; KRAB; 1.
CC SMART; SM00431; LER; 1.
CC SMART; SM00355; Znf_C2H2; 7.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 7.
CC Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 38 132 SCAN BOX.
FT ZN_FING 326 348 C2H2-TYPE.
FT ZN_FING 354 376 C2H2-TYPE.
FT ZN_FING 382 404 C2H2-TYPE.
FT ZN_FING 437 459 C2H2-TYPE.
FT ZN_FING 465 487 C2H2-TYPE.
FT ZN_FING 493 515 C2H2-TYPE.
SQ SEQUENCE 520 AA; 59155 MW; E9EDE71B85D1434A CRC64;
Query Match 71.8%; Score 44.5; DB 1; Length 520;
Best Local Similarity 42.9%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
QY 1 CPEC-----XHQTH 10
Db 384 CPEGCECFQSSHLRQTH 404
RESULT 7
ZF94_RAT
ID ZF94_RAT PRT; 536 AA.
AC Q92K3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 94 (Zfp-94) (zinc finger protein Y1) (RLZF-Y).
 GN ZFP94.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE=lung;
 RX MEDLINE=99023763; PubMed=9804994;
 RA Dovat S., Gilbert K.A., Petrovic-Dovat L., Rannels D.E.;
 RT "Isolation, cloning, and characterization of a novel rat lung zinc
 RT finger gene, RLZF-Y";
 RL Biochim. Biophys. Acta 1442:380-388(1998).
 CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -!- SIMILARITY: Contains 1 SCAN box domain.
 CC -!- SIMILARITY: Contains 7 C2H2-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF052042; AAC78780.1; -;
 CC HSSP; P08047; 1SP1.
 DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR003309; Treg_SCAN.
 DR InterPro: IPR007087; Znf_C2H2.
 DR InterPro: IPR007086; Znf_C2H2_sub.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF02023; SCAN; 1.
 DR Pfam; PF00096; zf-C2H2; 7.
 DR PRINTS; PR00048; ZINC_FINGER.
 DR ProDom; PD000003; Znf_C2H2; 6.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00431; LER; 1.
 DR SMART; SM00355; Znf_C2H2; 7.
 DR PROSITE; PS00804; SCAN_BOX; 1.
 DR PROSITE; P500028; ZINC_FINGER_C2H2_1; 6.
 DR PROSITE; P50157; ZINC_FINGER_C2H2_2; 7.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat.
 KW DOMAIN 44 126 SCAN BOX.
 FT ZN_FING 328 350 C2H2-TYPE 1.
 FT ZN_FING 356 378 C2H2-TYPE 2.
 FT ZN_FING 384 406 C2H2-TYPE 3.
 FT ZN_FING 412 433 C2H2-TYPE 4 (ATYPICAL).
 FT ZN_FING 439 461 C2H2-TYPE 5.
 FT ZN_FING 467 489 C2H2-TYPE 6.
 FT ZN_FING 495 517 C2H2-TYPE 7.
 SQ SEQUENCE 536 AA; 60987 MW; 9359894B2565ACD4 CRC64;
 Query Match 71.8%; Score 44.5; DB 1; Length 536;
 Best Local Similarity 42.9%; Pred. No. 1.2;
 Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
 QY 1 CPECGECFRRSSLSRQRTH 10
 Db 386 CPECGECFRRSSLSRQRTH 406
 RESULT 8
 ZN42_HUMAN
 ID ZN42_HUMAN STANDARD; PRT; 734 AA.
 AC P28698; O9UBW2;
 DT 01-DEC-1992 (Rel. 24, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Zinc finger protein 42 (Myeloid zinc finger 1) (MZFL-1).
 GN ZNF42 OR MZF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM MZF1A).
 RX MEDLINE=91317761; PubMed=1860835;
 RA Hromas R., Collins S.J., Hickstein D., Raskind W., Deaven L.L.,
 RA O'Hara P., Hagen F.S., Kaushansky K.;
 RT "A retinoic acid-responsive human zinc finger gene, MZF-1,
 RT preferentially expressed in myeloid cells";
 RL J. Biol. Chem. 266:14183-14187(1991).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM MZF1B-C).
 RC TISSUE=Bone marrow;
 RX MEDLINE=20432092; PubMed=10974541;
 RA Peterson M.J., Morris J.F.;
 RT "Human myeloid zinc finger gene MZF produces multiple transcripts and
 RT encodes a SCAN box protein.";
 RL Gene 254:105-118(2000).
 CC -!- FUNCTION: MAY BE ONE REGULATOR OF TRANSCRIPTIONAL EVENTS DURING
 CC HEMOPOIETIC DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=MZF1A;
 CC IsoId=P28698-1; Sequence=Displayed;
 CC Name=MZF1B-C;
 CC IsoId=P28698-2; Sequence=VSP_006889, VSP_006890;
 CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN DIFFERENTIATING
 CC MYELOID CELLS.
 CC -!- INDUCTION: By retinoic acid.
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -!- SIMILARITY: Contains 1 SCAN box domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M58297; AAA59898.1; -;
 CC EMBL; AF055078; AAD55810.1; -;
 CC EMBL; AF055077; AAD55809.1; -;
 CC EMBL; AF161886; AAF80465.1; -;
 CC HSSP; P08047; 1SP2.
 DR TRANSFAC; T00529; -;
 DR Genew; HGNC:13108; ZNF42.
 DR MIM; 194530; -;
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
 DR InterPro: IPR003309; Treg_SCAN.
 DR InterPro: IPR007087; Znf_C2H2.
 DR InterPro: IPR007086; Znf_C2H2_sub.
 DR Pfam; PF02023; SCAN; 1.
 DR Pfam; PF00096; zf-C2H2; 13.
 DR PRINTS; PR00048; ZINC_FINGER.
 DR ProDom; PD000003; Znf_C2H2; 6.
 DR SMART; SM00431; LER; 1.
 DR SMART; SM00355; Znf_C2H2; 13.
 DR PROSITE; PS00804; SCAN_BOX; 1.
 DR PROSITE; P500028; ZINC_FINGER_C2H2_1; 13.
 DR PROSITE; P50157; ZINC_FINGER_C2H2_2; 13.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat; Alternative splicing; Polymorphism.
 FT DOMAIN 44 125 SCAN BOX.
 FT DOMAIN 310 321 ASP/GLU-RICH (ACIDIC).

```

FT 2N_FING 356 378 C2H2-TYPE.
FT 2N_FING 384 406 C2H2-TYPE.
FT 2N_FING 412 434 C2H2-TYPE.
FT 2N_FING 440 462 C2H2-TYPE.
FT 2N_FING 463 484 GLY/PRO-RICH.
FT 2N_FING 485 507 C2H2-TYPE.
FT 2N_FING 513 535 C2H2-TYPE.
FT 2N_FING 541 563 C2H2-TYPE.
FT 2N_FING 569 591 C2H2-TYPE.
FT 2N_FING 597 619 C2H2-TYPE.
FT 2N_FING 625 647 C2H2-TYPE.
FT 2N_FING 653 675 C2H2-TYPE.
FT 2N_FING 681 703 C2H2-TYPE.
FT 2N_FING 709 731 C2H2-TYPE.
FT VARSPLIC 1 249 Missing (in isoform MZFlB-C).
FT VARSPLIC 250 257 /FTid-VSP_006889.
FT VARSPLIC 250 257 EAGGIFSP -> MNGPLVYA (in isoform MZFlB-C).
FT VARSPLIC 250 257 /FTid-VSP_006890.
FT VARSPLIC 250 257 I -> V (in dbSNP:4756).
FT VARSPLIC 250 257 /FTid-VAR_014826.
FT VARSPLIC 250 257 AL -> RV (in REF: 1).
FT VARSPLIC 250 257 734 AA; 82036 MW; 2BE7D69B18F29437 CRC64;
SQ SEQUENCE 734 AA; 82036 MW; 2BE7D69B18F29437 CRC64;

Query Match 71.8%; Score 44.5; DB 1; Length 734;
Best Local Similarity 42.9%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

Qy 1 CPEC-----XHQTH 10
Db 599 CPEGGRFSQRLKLTQH 619
|||||
|||||
599 CPEGGRFSQRLKLTQH 619

RESULT 9
XFIN_XENLA STANDARD; PRT: 1350 AA.
AC P08045;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein Xfin.
GN XFIN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=88082679; PubMed=2826129;
RA Ruiz i Altaba A., Perry-O'Keefe H., Melton D.A.;
RT *Xfin: an embryonic gene encoding a multifingered protein in
RT xenopus.*;
RL EMBO J. 6:3065-3070(1987).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94021366; PubMed=7692399;
RA Andreazoli M., de Lucchini S., Costa M., Barsacchi G.;
RT *RNA binding properties and evolutionary conservation of the Xenopus
RT multifinger protein Xfin.*;
RL Nucleic Acids Res. 21:4218-4225(1993).
RN [3]
RP STRUCTURE BY NMR OF FINGER 31.
RX MEDLINE=89346749; PubMed=2503871;
RA Lee M.S., Gippert G.P., Sonan K.V., Case D.A., Wright P.E.;
RT *Three-dimensional solution structure of a single zinc finger DNA-
RT binding domain.*;
RL Science 245:635-637(1989).
RN [4]
RP STRUCTURE BY NMR OF A FINGER.
RX MEDLINE=89378224; PubMed=2506074;
RA Lee M.S., Cavanagh J., Wright P.E.;

```

```

RT "Complete assignment of the 1H NMR spectrum of a synthetic zinc
RT finger from Xfin. Sequential resonance assignments and secondary
RT structure.";
RL FEBS Lett. 254:159-164(1989).
CC -!- FUNCTION: BINDS RNA. COULD FUNCTION IN POST-TRANSLATIONAL
CC REGULATION PROCESSES.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: OOCYTES AND IN SPECIALIZED CELL TYPES SUCH
CC AS NEURAL RETINA CONES.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT OOOGENESIS AND
CC EMBRYOGENESIS.
CC -!- DOMAIN: CONTAINS 37 FINGER MOTIFS IN 6 DOMAINS.
CC -!- PTM: PHOSPHORYLATED.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X06021; CAA29425.1; -.
DR PDB; 1ZNF; 15-OCT-91.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 36.
DR ProDom; PD000003; Znf_C2H2; 20.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 35.
DR PROSITE; PS0805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 35.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 37.
KW Zinc-finger; Metal-binding; RNA-binding; Repeat; 3D-structure;
KW Phosphorylation.
FT DOMAIN 1 58 KRAB.
FT 2N_FING 108 130 C2H2-TYPE.
FT 2N_FING 136 158 C2H2-TYPE.
FT 2N_FING 164 186 C2H2-TYPE.
FT 2N_FING 192 214 C2H2-TYPE.
FT 2N_FING 220 242 C2H2-TYPE.
FT 2N_FING 248 270 C2H2-TYPE.
FT 2N_FING 276 298 C2H2-TYPE.
FT 2N_FING 326 348 C2H2-TYPE.
FT 2N_FING 354 376 C2H2-TYPE.
FT 2N_FING 382 404 C2H2-TYPE.
FT 2N_FING 410 432 C2H2-TYPE.
FT 2N_FING 438 460 C2H2-TYPE.
FT 2N_FING 466 488 C2H2-TYPE.
FT 2N_FING 503 525 C2H2-TYPE.
FT 2N_FING 531 553 C2H2-TYPE.
FT 2N_FING 559 581 C2H2-TYPE.
FT 2N_FING 587 609 C2H2-TYPE.
FT 2N_FING 615 637 C2H2-TYPE.
FT 2N_FING 643 665 C2H2-TYPE.
FT 2N_FING 671 693 C2H2-TYPE.
FT 2N_FING 699 721 C2H2-TYPE.
FT 2N_FING 750 772 C2H2-TYPE.
FT 2N_FING 778 800 C2H2-TYPE.
FT 2N_FING 806 828 C2H2-TYPE.
FT 2N_FING 834 856 C2H2-TYPE.
FT 2N_FING 862 884 C2H2-TYPE.
FT 2N_FING 890 912 C2H2-TYPE.
FT 2N_FING 918 940 C2H2-TYPE.
FT 2N_FING 988 1010 C2H2-TYPE.
FT 2N_FING 1016 1038 C2H2-TYPE.
FT 2N_FING 1044 1066 C2H2-TYPE.
FT 2N_FING 1136 1158 C2H2-TYPE.
FT 2N_FING 1164 1186 C2H2-TYPE.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

DR EMBL; AK001252; BAA91582.1; -
DR EMBL; BC003585; AAH03585.1; -
DR EMBL; BC014002; AAH14002.1; -
DR HSSP; P25490; IUBD
DR Genew; HGNC:16838; ZNF358.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 9.
DR ProDom; PD000003; Znf_C2H2; 4.
DR SMART; PS00028; ZINC_FINGER_C2H2_1; 9.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 9.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 9.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT ZN_FING 64 86 C2H2-TYPE.
FT ZN_FING 92 114 C2H2-TYPE.
FT ZN_FING 120 142 C2H2-TYPE.
FT ZN_FING 148 170 C2H2-TYPE.
FT ZN_FING 176 198 C2H2-TYPE.
FT ZN_FING 204 226 C2H2-TYPE.
FT ZN_FING 232 254 C2H2-TYPE.
FT ZN_FING 260 282 C2H2-TYPE.
FT ZN_FING 288 310 C2H2-TYPE.
FT DOMAIN 313 329 POLY-ALA.
SQ SEQUENCE 481 AA; 49656 MW; 93EDB718A799BC0E CRC64;

Query Match 66.9%; Score 41.5; DB 1; Length 481;
Best Local Similarity 38.1%; Pred. No. 3.5;
Matches 8; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
DB 206 CPQCGKFGQSALLQHQTH 226
|||||

RESULT 11
2382_HUMAN
ID 2382_HUMAN STANDARD; PRT; 548 AA.
AC Q96SR6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 382.
GN ZNF382.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Luo K., Liu M., Wu X.;
RT "Cloning and characterization of a novel Kruppel-like zinc finger gene, ZNF382, specially expressed in adult human heart.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NED0 human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May function as a transcription factor.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in heart.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.

CC FINGER PROTEINS.
CC -1- SIMILARITY: Contains 1 KRAB domain.
CC -1- SIMILARITY: Contains 10 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF513816; RA048246.1; -
CC EMBL: AK027592; BAB55217.1; -
CC InterPro: IPR001909; KRAB.
CC InterPro: IPR007087; Znf_C2H2.
CC Pfam: PF01352; KRAB; 1.
CC Pfam: PF00096; zf-C2H2; 9.
CC ProDom: PD000003; Znf_C2H2; 3.
CC PROSITE: PS00805; KRAB; 1.
CC PROSITE: PS00028; ZINC_FINGER_C2H2_1; 9.
CC PROSITE: PS0157; ZINC_FINGER_C2H2_2; 10.
CC Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;
CC DNA-binding; Repeat.
CC DOMAIN 7 KRAB.
CC FT ZN_FING 210 232 C2H2-TYPE 1 (DEGENERATE).
CC FT ZN_FING 294 316 C2H2-TYPE 2.
CC FT ZN_FING 322 344 C2H2-TYPE 3.
CC FT ZN_FING 372 392 C2H2-TYPE 4.
CC FT ZN_FING 358 400 C2H2-TYPE 5.
CC FT ZN_FING 406 428 C2H2-TYPE 6.
CC FT ZN_FING 434 456 C2H2-TYPE 7.
CC FT ZN_FING 462 484 C2H2-TYPE 8.
CC FT ZN_FING 490 512 C2H2-TYPE 9.
CC FT ZN_FING 518 540 C2H2-TYPE 10.
CC SEQUENCE 548 AA: 63794 MW: C25B782CB253CC9 CRC64;

Query Match 66.9%; Score 41.5; DB 1; Length 548;
Best Local Similarity 38.1%; Pred. No. 4;
Matches 8; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XKORTH 10
||:|
DB 380 CPGCGAFKRSYLDHQRIH 400
|||||

RESULT 12
ID 2263 HUMAN STANDARD; PRT; 683 AA.
AC 014978; 043387; Q96H95;
DT 30-MAY-2000 (Rel. 39, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 263 (Zinc finger protein FPM315).
GN ZNF263 OR FPM315.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97398134; PubMed-9256059;
RA Yokoyama M., Nakamura M., Okudo K., Matsubara K., Nishi Y.,
RA Matsumoto T., Fukushima A.;
RT Isolation of a cDNA encoding a widely expressed novel zinc finger
RT protein with the LeR and KRAB-A domains.*;
RL Biochim. Biophys. Acta 1353:13-17(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Ricke D.O., Bruce D., Mundt M., Doggett N., Munk C., Saunders E.,
RA Robinson D., Jones M., Buckingham J., Chasteen L., Thompson S.,
RA Goodwin L., Bryant J., Tesmer J., Meinkne L., Longmire J., White S.,
RA Deng S., Tatum O., Campbell C., Fawcett J., Maltbie M., Mista M.,

RA Deaven L.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE-22388257; PubMed-12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RX Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RX Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RX Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RX Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RX Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RX Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RX Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RX Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guntaratne P.H.,
RX Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RX Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RX Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RX Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RX Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RX Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Smailus D.E.,
RX Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT *Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.*;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: MIGHT PLAY AN IMPORTANT ROLE IN BASIC CELLULAR
CC PROCESSES AS A TRANSCRIPTIONAL REPRESSOR.
CC CC -1- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,
CC LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS, SPLEEN, THYMUS,
CC PROSTATE, TESTIS, OVARY, SMALL INTESTINE, COLON AND LEUKOCYTE.
CC CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC CC -1- SIMILARITY: Contains 1 SCAN box domain.
CC CC -1- SIMILARITY: Contains 1 KRAB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D88827; BAA21853.1; -
CC EMBL: AC004232; AAC24490.1; -
CC EMBL: BC008805; AAH08805.1; -
CC HSP: P08045; IZNF.
CC Genew: HGNC:13056; ZNF263.
CC MIM: 604191; -
CC GO: GO:0003700; F:transcription factor activity; TAS.
CC GO: GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
CC InterPro: IPR001909; KRAB.
CC InterPro: IPR003309; Treg_SCAN.
CC InterPro: IPR007087; Znf_C2H2.
CC InterPro: IPR007086; Znf_C2H2_sub.
CC Pfam: PF01352; KRAB; 1.
CC Pfam: PF02023; SCAN; 1.
CC Pfam: PF00096; zf-C2H2; 9.
CC PRINTS: PR00048; ZINC_FINGER.
CC ProDom: PD000003; Znf_C2H2; 9.
CC SMART: SM00349; KRAB; 1.
CC SMART: SM00431; LER; 1.
CC SMART: SM00355; Znf_C2H2; 9.
CC PROSITE: PS00805; KRAB; 1.
CC PROSITE: PS00028; ZINC_FINGER_C2H2_1; 9.
CC PROSITE: PS0157; ZINC_FINGER_C2H2_2; 9.
CC Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;
CC DNA-binding; Repeat; Repressor.
FT DOMAIN 41 123 SCAN BOX.
FT DOMAIN 217 289 KRAB.

FT DOMAIN 153 191 RUBREDOXIN-LIKE.
FT METAL 20 20 ZINC OR IRON 1.
FT METAL 53 53 ZINC OR IRON 1.
FT METAL 53 53 IRON 2.
FT METAL 94 94 IRON 2.
FT METAL 97 97 ZINC OR IRON 1.
FT METAL 128 128 ZINC OR IRON 1.
FT METAL 128 128 IRON 2.
FT METAL 131 131 IRON 2.
FT METAL 158 158 IRON 3.
FT METAL 161 161 IRON 3.
FT METAL 174 174 IRON 3.
FT METAL 177 177 IRON 3.
FT METAL 177 177 IRON 3.
FT HELIX 8 37
FT TURN 38 38
FT TURN 40 61
FT HELIX 40 61
FT TURN 62 63
FT HELIX 83 99
FT TURN 100 100
FT HELIX 101 111
FT TURN 112 113
FT HELIX 115 143
FT TURN 144 145
FT STRAND 149 158
FT TURN 159 161
FT STRAND 164 169
FT STRAND 173 173
FT TURN 175 177
FT STRAND 180 180
FT HELIX 181 183
FT STRAND 184 186
SQ SEQUENCE 191 AA: 21544 MW: BCF8E9A2659F1138 CRC64;

Query Match 66.1%; Score 41; DB 1; Length 191;
Best Local Similarity 50.0%; Pred. No. 1.8;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CPECHXQHTH 10
||| | | |
DB 174 CPACAPRAH 183

RESULT 14
CHS6_YEAST STANDARD; PRT: 746 AA.
AC P40955;
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Chitin biosynthesis protein CHS6 (CSB3 protein).
GN CHS6 OR CSB3 OR YJL099W OR J0838.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Osmond B.C.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=S288c / FY1679;
RX MEDLINE=96090136; PubMed=7483851;
RA Rasmussen S.W.;
RT A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1
RT and CSB3 genes, a TCP-1-related gene, an open reading frame similar
RT to the DAL80 gene, and a tRNA(Arg).
RL Yeast 11:873-883(1995).
CC -!- FUNCTION: NOT KNOWN: CSB3 MUTANTS LACK CHITIN IN VIVO.
CC -!- SIMILARITY: TO YEAST YKR027W AND YMR237W.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: U15603; AAA50840.1; -
DR EMBL: X85021; CAA59395.1; -
DR EMBL: Z49374; CAA89394.1; -
DR EMBL: Z49373; CAA89393.1; -
DR PIR: S50226; S50226.
DR SGD: S0003635; CHS6.
DR GO: GO:0006038; P:cell wall chitin biosynthesis; IMP.
DR GO: GO:0008893; P:Golgi to plasma membrane transport; IMP.
DR InterPro: IPR001440; TPR.
SQ SEQUENCE 746 AA; 86042 MW; BD92DADA3D5396BC CRC64;

Query Match 66.1%; Score 41; DB 1; Length 746;
Best Local Similarity 44.4%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

OY 1 CP-----ECXHORTH 10
||| | | | |
DB 43 CPRIMERQFGECLHNRTH 60

RESULT 15
ZF13_MOUSE STANDARD; PRT: 82 AA.
AC P10754;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 13 (Zfp-13) (Krox-8 protein) (Fragment).
GN ZFP13 OR ZFP-13 OR KROX-8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88216608; PubMed=2452975;
RA Chavrier P., Lemaire P., Revelant O., Bravo R., Charnay P.;
RT "Characterization of a mouse multigene family that encodes zinc
RT finger structures";
RL Mol. Cell. Biol. 8:1319-1326(1988).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: M19440; AAA39388.1; -
DR PIR: E29942; E29942.
DR HSP: P08046; 1ALC.
DR MGD: MGI:99159; Zfp13.
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam: PF000096; zf-C2H2; 3.
DR ProDom: PD000003; Znf_C2H2; 2.
DR SMART: SM00355; Znf_C2H2; 3.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 3.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 3.
KW Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat.
FT NON_TER 1 1
FT ZN_FING 7 29 C2H2-TYPE.
FT ZN_FING 35 57 C2H2-TYPE.
FT ZN_FING 63 >82 C2H2-TYPE.
FT NON_TER 82 82
SQ SEQUENCE 82 AA; 9350 MW; 51349DE47199C6E8 CRC64;

Query Match 65.3%; Score 40.5; DB 1; Length 82;
Best/Local Similarity 38.1%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 2; Indels 11; Gaps 1;

QY 1 CPEC-----XHQETH 10
| | | |
| | | |
Db 9 CPEGKCFQSRSLIAHNRTH 29

Search completed: October 1, 2003, 10:42:36
Job time : 12.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:33:32 ; Search time 59.7222 Seconds
(without alignments)
43.209 Million cell updates/sec

Title: US-10-057-890A-3
Perfect score: 62
Sequence: 1 CPEXHQTH 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_unclassified:*
14: sp_viruses:*
15: sp_bacteriaph:*
16: sp_bacteriaph:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	45.5	73.4	720	6 Q9N003	Q9N003 macaca fasc
2	45.5	73.4	812	11 O35483	O35483 mus musculus
3	45.5	73.4	841	11 Q9WV18	Q9WV18 mus musculus
4	44.5	71.8	38	11 P97651	P97651 rattus norv
5	44.5	71.8	90	4 O14892	O14892 homo sapien
6	44.5	71.8	105	6 Q8MIE8	Q8MIE8 globicephal
7	44.5	71.8	105	6 Q8MIA0	Q8MIA0 sus scrofa
8	44.5	71.8	105	6 Q8MIO0	Q8MIO0 cebuella py
9	44.5	71.8	203	11 Q8BIN3	Q8BIN3 mus musculus
10	44.5	71.8	208	4 Q8N428	Q8N428 homo sapien
11	44.5	71.8	223	11 Q8BIN4	Q8BIN4 mus musculus
12	44.5	71.8	230	4 Q8BTG0	Q8BTG0 homo sapien
13	44.5	71.8	315	11 Q9CZG8	Q9CZG8 mus musculus
14	44.5	71.8	318	11 Q9D9Y1	Q9D9Y1 mus musculus
15	44.5	71.8	353	11 Q8R0W4	Q8R0W4 mus musculus
16	44.5	71.8	452	4 Q9H6G5	Q9H6G5 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9N003	PRELIMINARY:	PRT:	720 AA.
AC	Q9N003;			
DT	01-OCT-2000 (TRENBLrel. 15, Created)			
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)			
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)			
DE	Hypothetical 83.9 kDa protein.			
OS	Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;			
OC	Cercopitheidae; Macaca.			
OX	NCBI_TaxID=9541;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Cerebellum;			
RA	Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Tetao K.,			
RA	Suzuki Y., Sugano S., Hashimoto K.			
RT	*Isolation of full-length cDNA clones from macaque brain cDNA			
RT	libraries.*;			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB046644; BAB03562.1; -			
DR	HSSP; P07248; LARE.			
DR	InterPro; IPR001909; KRAB.			
DR	InterPro; IPR007087; Znf_C2H2.			
DR	Pfam; PF000096; zf-C2H2; 19.			
DR	ProDom; PD000003; Znf_C2H2; 2.			
DR	SMART; SM00349; KRAB; 1.			
DR	SMART; SM00355; Znf_C2H2; 19.			
DR	PROSITE; PS00805; KRAB; 1.			
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 19.			
DR	PROSITE; PS0157; ZINC_FINGER_C2H2_2; 19.			
KW	Hypothetical protein; Metal-binding; Zinc; zinc-finger.			
SQ	SEQUENCE 720 AA; 83922 MW; 074A7736A5903D90 CRC64;			

Query Match 73.4%; Score 45.5; DB 6; Length 720;
Best Local Similarity 42.9%; Pred.No.3.7;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

Q96CQ8 homo sapien
Q99K45 mus musculus
Q8B111 mus musculus
Q99LH4 mus musculus
Q8B1R7 mus musculus
Q9NRY0 homo sapien
Q60304 homo sapien
Q9CYV9 mus musculus
Q9UQ14 homo sapien
Q96EG3 homo sapien
Q9Y2P0 homo sapien
Q8TC91 homo sapien
Q8K379 mus musculus
Q8CGF0 mus musculus
Q96I71 homo sapien
Q84N4 mus musculus
Q70162 mus musculus
Q8J18 infectious
Q82739 infectious
Q8IY24 homo sapien
Q8B173 mus musculus
Q96BV0 homo sapien
P97644 rattus norv
Q15938 homo sapien
Q8SYC1 drosophila
Q923T4 mus musculus
Q8N1Q0 homo sapien
Q9HAQ4 homo sapien
Q96G73 homo sapien

17 44.5 71.8 454 4 Q96CQ8
18 44.5 71.8 461 11 Q99K45
19 44.5 71.8 461 11 Q8B111
20 44.5 71.8 468 11 Q99LH4
21 44.5 71.8 468 11 Q8B1R7
22 44.5 71.8 485 4 Q9NRY0
23 44.5 71.8 493 4 Q60304
24 44.5 71.8 521 11 Q9CYV9
25 44.5 71.8 540 4 Q9UQ14
26 44.5 71.8 558 4 Q96EG3
27 44.5 71.8 559 4 Q9Y2P0
28 44.5 71.8 573 4 Q8TC91
29 44.5 71.8 573 11 Q8K379
30 44.5 71.8 679 11 Q8CGF0
31 44.5 71.8 734 4 Q96I71
32 44.5 71.8 735 11 Q84N4
33 44.5 71.8 814 11 Q70162
34 43 69.4 147 12 Q8J18
35 43 69.4 147 12 Q82739
36 42.5 68.5 372 4 Q8IY24
37 42.5 68.5 538 11 Q8B173
38 42.5 68.5 561 4 Q96BV0
39 41.5 66.9 36 11 P97644
40 41.5 66.9 109 4 Q15938
41 41.5 66.9 157 5 Q8SYC1
42 41.5 66.9 484 11 Q923T4
43 41.5 66.9 529 4 Q8N1Q0
44 41.5 66.9 532 4 Q9HAQ4
45 41.5 66.9 532 4 Q96G73

QY 1 CPEC-----XHQTH 10
|||||
328 CPEGKRSFQRKALKTHQTH 348

RESULT 2
O35483 PRELIMINARY: PRT; 812 AA.
ID O35483
AC O35483;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Kruppel-related zinc finger protein.
GN ZFP316 OR EMZF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Volk sac;
RA Rohsstein J.L., Brafford P.;
RT "Murine Emzfl gene: Expression During Development.";
RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031955; AAB87452.1; -;
DR HSSP; P08045; 12NF.
DR MGD; MGI:1860402; Zfp316.
DR InterPro; IPR000515; BPD.transp.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 14.
DR ProDom; PD000003; Znf_C2H2; 6.
DR SMART; SM00355; Znf_C2H2; 14.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 14.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 812 AA; 88617 MW; 8F9DDE55FD063DCB CRC64;

Query Match 73.4%; Score 45.5; DB 11; Length 812;
Best Local Similarity 42.9%; Pred. No. 4.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
637 CPEGKRSFQRSVLVTHQTH 657

RESULT 3
O9WVL8 PRELIMINARY: PRT; 841 AA.
ID O9WVL8
AC O9WVL8;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Zinc finger type transcription factor MZF-3.
GN ZFP316.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B57BL/6;
RA Yamaguchi Y., Zhang X.O., Suda T.;
RT "Cloning of MZF-1 related gene, murine MZF-3.";
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF082568; AAD38426.1; -;
DR HSSP; P08045; 12NF.
DR MGD; MGI:1860402; Zfp316.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 15.
DR ProDom; PD000003; Znf_C2H2; 7.
DR SMART; SM00355; Znf_C2H2; 15.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 15.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 841 AA; 91149 MW; 8AB1E4D2A02C3516 CRC64;

Query Match 73.4%; Score 45.5; DB 11; Length 841;
Best Local Similarity 42.9%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
730 CPEGKRSFQRSVLVTHQTH 750

RESULT 4
P97651 PRELIMINARY: PRT; 38 AA.
ID P97651
AC P97651;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Zinc finger protein 12 (Fragment).
GN AZF12.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96220733; PubMed=8635150;
RA Mesa K., Gebelein B., Cook T., Urrutia R.;
RT "Identification and characterization of zinc finger encoding genes
from the tumoral exocrine pancreatic cell line AR42J.";
RL Cancer Lett. 103:143-149(1996).
DR EMBL; U78123; AAB36795.1; -;
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 38
SQ SEQUENCE 38 AA; 4312 MW; 3B1C3B42B8F410E CRC64;

Query Match 71.8%; Score 44.5; DB 11; Length 38;
Best Local Similarity 42.9%; Pred. No. 0.46;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
2 CPEGKRSFQRSVLVTHQTH 22

RESULT 5
O14892 PRELIMINARY: PRT; 90 AA.
ID O14892
AC O14892;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Zinc finger protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-72 FROM N.A.
RC TISSUE=Brain;
RA Sun X., Yu L., Wu G., Liu S., Hu P., Zhang M., Jiang Y., Shouyuan S.;
RT "Isolation of novel human genes coding zinc finger protein from brain
tissue - according to the conservativity of zinc finger motif.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 0:0-0(1997).

RN
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RA Zhang Q.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF027147; AAB84026.1; -.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF000096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Metal-binding; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 10312 MW; 5A5F670BA6D37C47 CRC64;

Query Match 71.8%; Score 44.5; DB 4; Length 90;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
DB 1 CPEGKSFQKVTSTQHQTH 21
|||||

RESULT 6
Q8MIE8
ID Q8MIE8 PRELIMINARY; PRT; 105 AA.
AC Q8MIE8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Zinc finger protein 33b (Fragment).
GN ZNF33B.
OS Globicephala macrorhynchus (short-finned pilot whale).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC Globicephala.
OX NCBI_TaxID=38241;
RN
RP SEQUENCE FROM N.A.
RA Hearn T.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RP
RP SEQUENCE FROM N.A.
RA Guy J., Hearn T., Crosier M., Mudge J., Viggiano L., Koczan D.,
RA Thiesen H.J., Bailey J., Horvath J., Eichler E.E., Deloukas P.,
RA French L., Rogers J., Bentley D., Jackson M.S.;
*Genomic sequence and transcriptional profile of the boundary between
RT pericentromeric satellites and genes on human chromosome arm 10p.*;
RL Genome Res. 0:0-0(0).
RN
RP
RP SEQUENCE FROM N.A.
RA Hearn T.;
RA *Organisation, expression and evolution of Kruppel-type zinc finger
RT genes in human chromosomal region 10p11.2-q11.2*;
RL Thesis (2000), Department of Human Genetics, University of Newcastle,
RL Newcastle upon Tyne, United Kingdom.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AJ491694; CA036953.1;
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 12174 MW; 0C27F798AAC9D7E1 CRC64;

Query Match 71.8%; Score 44.5; DB 6; Length 105;
Best Local Similarity 42.9%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
DB 69 CPEGKSFQKSHLTQHQTH 89
|||||

RESULT 7
Q8MIAO
ID Q8MIAO PRELIMINARY; PRT; 105 AA.
AC Q8MIAO;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Zinc finger protein 33b (Fragment).
GN ZNF33B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN
RP SEQUENCE FROM N.A.
RA Hearn T.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RP
RP SEQUENCE FROM N.A.
RA Guy J., Hearn T., Crosier M., Mudge J., Viggiano L., Koczan D.,
RA Thiesen H.J., Bailey J., Horvath J., Eichler E.E., Deloukas P.,
RA French L., Rogers J., Bentley D., Jackson M.S.;
*Genomic sequence and transcriptional profile of the boundary between
RT pericentromeric satellites and genes on human chromosome arm 10p.*;
RL Genome Res. 0:0-0(0).
RN
RP
RP SEQUENCE FROM N.A.
RA Hearn T.;
RA *Organisation, expression and evolution of Kruppel-type zinc finger
RT genes in human chromosomal region 10p11.2-q11.2*;
RL Thesis (2000), Department of Human Genetics, University of Newcastle,
RL Newcastle upon Tyne, United Kingdom.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AJ491691; CAD36950.1; -.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 12210 MW; 21F3D1CC9AD15EEC CRC64;

Query Match 71.8%; Score 44.5; DB 6; Length 105;
Best Local Similarity 42.9%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
DB 69 CPEGKSFQKSHLTQHQTH 89
|||||

RESULT 8
Q8MIHO
ID Q8MIHO PRELIMINARY; PRT; 105 AA.
AC Q8MIHO;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Zinc finger protein 33b (Fragment).
GN ZNF33B.
OS Cebueilla pygmaea (Pygmy marmoset) (Callithrix pygmaea).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
OX NCBI_TaxID=9493;
RN [1]
RP SEQUENCE FROM N.A.
RA Hearn T.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Guy J., Hearn T., Crosier M., Mudge J., Viggiano L., Koczan D.,
RA Thiesen H.J., Bailey J., Horvath J., Eichler E.E., Deloukas P.,
RA French L., Rogers J., Bentley D., Jackson M.S.;
RT Genomic sequence and transcriptional profile of the boundary between
RT pericentromeric satellites and genes on human chromosome arm 10p.*;
RN Genome Res. 0:0-0(0).
RN [3]
RP SEQUENCE FROM N.A.
RA Hearn T.;
RT *Organisation, expression and evolution of Kruppel-type zinc finger
RT genes in human chromosomal region 10p11.2-q11.2.*;
RL Thesis (2000), Department of Human Genetics, University of Newcastle,
RL Newcastle upon Tyne, United Kingdom.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AJ491692; CAD36951.1; -.
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam: PF000096; Znf_C2H2; 3.
DR PRINTS: PR00048; ZINCFINGER.
DR PRODOM: PD000003; Znf_C2H2; 1.
DR SMART: SM00355; Znf_C2H2; 3.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 3.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 105
FT SEQUENCE 105 AA; 12130 MW; 494C9FDE7CADEA8 CRC64;

Query Match 71.8%; Score 44.5; DB 6; Length 105;
Best Local Similarity 42.9%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|
|
|
|
DB 69 CPEGKSFQKSHLTQHORTH 89

RESULT 9
Q8BIN3 PRELIMINARY; PRT; 203 AA.
ID Q8BIN3
AC Q8BIN3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Weakly similar to ZNF74-KRUPPEL-type zinc finger.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT *Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.*;
RL Nature 420:563-573(2002).
DR EMBL: AK039299; BAC30310.1; -.
SQ SEQUENCE 203 AA; 22611 MW; 09460C36F1685022 CRC64;

Query Match 71.8%; Score 44.5; DB 11; Length 203;
Best Local Similarity 42.9%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
QY 1 CPEC-----XHQTH 10
|
|
|
|
DB 70 CPEGKAFSVKHNLEVHORTH 90
RESULT 10
Q8N4Z8 PRELIMINARY; PRT; 208 AA.
ID Q8N4Z8
AC Q8N4Z8
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to KIAA0557 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: BC033151; AAH3151.1; -.
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam: PF000096; Znf_C2H2; 5.
DR PRINTS: PR00048; ZINCFINGER.
DR SMART: SM00355; Znf_C2H2; 5.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 5.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 5.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 208 AA; 22892 MW; E9DDB492184A8BA3 CRC64;

Query Match 71.8%; Score 44.5; DB 4; Length 208;
Best Local Similarity 42.9%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|
|
|
|
DB 55 CPEGKGFSTSLTKHQTH 75

RESULT 11
Q8BIN4 PRELIMINARY; PRT; 223 AA.
ID Q8BIN4
AC Q8BIN4
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Weakly similar to ZNF74-KRUPPEL-type zinc finger.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT *Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.*;
RL Nature 420:563-573(2002).
DR EMBL: AK038960; BAC30184.1; -.
SQ SEQUENCE 223 AA; 23594 MW; C9EA5CC9A0CD3D01 CRC64;

Query Match 71.8%; Score 44.5; DB 11; Length 223;
Best Local Similarity 42.9%; Pred. No. 2.1;

```

Matches 9: Conservative 0: Mismatches 1: Indels 11: Gaps 1:
QY 1 CPEC-----XHQTH 10
    |||||
Db 116 CPEGKAFSVKHNLEHVH 136
    |||||

RESULT 12
Q9BTG0 PRELIMINARY: PRT; 230 AA.
AC Q9BTG0;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; BC004125; AAH04125.1; -.
DR HSSP; P08047; 1SP2.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf-C2H2; 5.
DR PRINTS; PR00048; ZNCFINGER.
DR SMART; SM00355; Znf_C2H2; 7.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 7.
KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
FT NON_TER
SQ SEQUENCE 230 AA; 25140 MW; 43A4E225BCC139F9 CRC64;

Query Match 71.8%; Score 44.5; DB 4; Length 230;
Best Local Similarity 42.9%; Pred. No. 2.1;
Matches 9: Conservative 0: Mismatches 1: Indels 11: Gaps 2:
QY 1 CPEC-----XHQTH 10
    |||||
Db 77 CPEGKGFSTSLTKHQTH 97
    |||||

RESULT 13
Q9CZG8 PRELIMINARY: PRT; 315 AA.
AC Q9CZG8;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE 2810437EL4RIK protein.
GN ZFP99 OR 2810437EL4RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bojunga M., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RA *Functional annotation of a full-length mouse cDNA collection.*;
RT Nature 409:685-690(2001).

```

```

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga M., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RA *Functional annotation of a full-length mouse cDNA collection.*;
RT Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AK012630; BAB28368.1; -.
DR HSSP; P08045; LZNF.
DR MGD; MGI:1914485; Zfp99.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf-C2H2; 7.
DR PRINTS; PR00048; ZNCFINGER.
DR PRODOM; PD000003; Znf_C2H2; 6.
DR SMART; SM00355; Znf_C2H2; 7.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 7.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ SEQUENCE 315 AA; 36116 MW; AA9453581B8B4A76 CRC64;

Query Match 71.8%; Score 44.5; DB 11; Length 315;
Best Local Similarity 42.9%; Pred. No. 2.8;
Matches 9: Conservative 0: Mismatches 1: Indels 11: Gaps 1:
QY 1 CPEC-----XHQTH 10
    |||||
Db 179 CPEGCEFRQSSHLRQH 199
    |||||

RESULT 14
Q9D9Y1 PRELIMINARY: PRT; 318 AA.
AC Q9D9Y1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE 1700025J14RIK protein.
GN 1700025J14RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bojunga M., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RA *Functional annotation of a full-length mouse cDNA collection.*;
RT Nature 409:685-690(2001).

```

DR EMBL; AK006349; BAB24541.1; -.
DR HSP; P08046; IAIG.
DR MGD; MGI:1915161; 1700025J14RIK.
DR InterPro; IPR007087; znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 11.
DR ProDom; PD000003; znf_C2H2; 8.
DR SMART; SM00355; znf_C2H2; 11.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 11.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 11.
KW Metal-binding; zinc; zinc-finger.
SQ SEQUENCE 318 AA; 37069 MW; DB34AF11DC9FAEDB CRC64;

Query Match 71.8%; Score 44.5; DB 11; Length 318;
Best Local Similarity 42.9%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 99 CPECGRFVKQSMILRHQTH 119
|||||

RESULT 15
Q8R0W4
ID Q8R0W4 PRELIMINARY; PRT; 353 AA.
AC Q8R0W4;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical 41.3 kda protein (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC026376; AAH26376.1; -.
DR InterPro; IPR007087; znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 10.
DR ProDom; PD000003; znf_C2H2; 2.
DR SMART; SM00355; znf_C2H2; 10.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 10.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 11.
KW Hypothetical protein; Metal-binding; zinc; zinc-finger.
FT NON_TER 1
SQ SEQUENCE 353 AA; 41330 MW; 1BE9BA53AE2661DF CRC64;

Query Match 71.8%; Score 44.5; DB 11; Length 353;
Best Local Similarity 42.9%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 136 CPECGLFYLKSTLTHQTH 156
|||||

Search completed: October 1, 2003, 10:46:21
Job time : 63.7222 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:42:07 ; Search time 26.6667 Seconds
(without alignments)
59.330 Million cell updates/sec

Title: US-10-057-890A-3
Perfect score: 62
Sequence: 1 CPECXHQTH 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	96.8	10	15	US-10-057-890A-3
2	45.5	73.4	98	15	US-10-113-424-10
3	45.5	73.4	100	9	US-09-989-789-15
4	45.5	73.4	100	11	US-09-846-0338-223
5	45.5	73.4	100	11	US-09-990-186-15
6	45.5	73.4	100	11	US-09-989-994-15
7	45.5	73.4	100	15	US-10-006-069A-223
8	45.5	73.4	184	11	US-09-765-555-38
9	45.5	73.4	184	11	US-09-765-555-39
10	45.5	73.4	184	11	US-09-765-555-75
11	45.5	73.4	186	10	US-09-908-153B-32
12	44.5	71.8	25	15	US-10-198-677-77
13	44.5	71.8	26	9	US-09-851-271A-8
14	44.5	71.8	26	11	US-09-995-973-8
15	44.5	71.8	26	11	US-09-996-484-8

16	44.5	71.8	28	10	US-09-908-153B-49	Sequence 49, Appl
17	44.5	71.8	28	10	US-09-908-153B-52	Sequence 52, Appl
18	44.5	71.8	28	12	US-09-911-261A-13	Sequence 13, Appl
19	44.5	71.8	28	12	US-09-911-261A-15	Sequence 15, Appl
20	44.5	71.8	28	12	US-09-911-261A-16	Sequence 16, Appl
21	44.5	71.8	28	12	US-09-911-261A-17	Sequence 17, Appl
22	44.5	71.8	28	12	US-09-911-261A-68	Sequence 68, Appl
23	44.5	71.8	28	12	US-09-911-261A-69	Sequence 69, Appl
24	44.5	71.8	28	15	US-10-055-713-11	Sequence 11, Appl
25	44.5	71.8	28	15	US-10-057-408-13	Sequence 13, Appl
26	44.5	71.8	28	15	US-10-057-408-15	Sequence 15, Appl
27	44.5	71.8	28	15	US-10-057-408-16	Sequence 16, Appl
28	44.5	71.8	28	15	US-10-057-408-17	Sequence 17, Appl
29	44.5	71.8	28	15	US-10-057-408-68	Sequence 68, Appl
30	44.5	71.8	28	15	US-10-057-408-69	Sequence 69, Appl
31	44.5	71.8	28	15	US-10-055-711-39	Sequence 19, Appl
32	44.5	71.8	34	10	US-09-908-153B-50	Sequence 50, Appl
33	44.5	71.8	34	10	US-09-908-153B-53	Sequence 53, Appl
34	44.5	71.8	38	10	US-09-908-153B-48	Sequence 48, Appl
35	44.5	71.8	38	10	US-09-908-153B-51	Sequence 51, Appl
36	44.5	71.8	84	12	US-09-911-261A-30	Sequence 30, Appl
37	44.5	71.8	84	15	US-10-057-408-30	Sequence 5, Appl
38	44.5	71.8	90	12	US-10-303-686A-5	Sequence 10, Appl
39	44.5	71.8	90	15	US-10-209-184-10	Sequence 5, Appl
40	44.5	71.8	90	15	US-10-147-286-5	Sequence 5, Appl
41	44.5	71.8	99	12	US-09-911-261A-5	Sequence 5, Appl
42	44.5	71.8	99	12	US-09-911-261A-6	Sequence 6, Appl
43	44.5	71.8	99	12	US-09-911-261A-7	Sequence 7, Appl
44	44.5	71.8	99	12	US-09-911-261A-8	Sequence 8, Appl
45	44.5	71.8	99	12	US-09-911-261A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-10-057-890A-3
; Sequence 3, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same
; FILE REFERENCE: PF537
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US/10/057,890A
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (5)..(5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids, one
; OTHER INFORMATION: residues may be missing
US-10-057-890A-3

Query Match 96.8% Score 60; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CPECXHQTH 10
| | | | | | | | | |
Db 1 CPECXHQTH 10

RESULT 2

```
US-10-113-424-10
; Sequence 10, Application US/10113424
; Publication No. US20030105593A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Case, Casey C.
; APPLICANT: Cox III, George N.
; APPLICANT: Jamieson, Andrew
; APPLICANT: Rebar, Edward J.
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
; TITLE OF INVENTION: Proteins and Methods of Designing Zinc Finger Proteins
; FILE REFERENCE: 019496-001800US
; CURRENT APPLICATION NUMBER: US/10/113,424
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US/09/229,007A
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Sp-1
; OTHER INFORMATION: transcription factor consensus sequence
US-10-113-424-10

Query Match 73.4%; Score 45.5; DB 15; Length 98;
Best Local Similarity 42.9%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
   ||||| |||||
Db 19 CPEGKSFSSKSHLRAHQTH 39

RESULT 3
US-09-989-789-15
; Sequence 15, Application US/09989789
; Patent No. US20020063379A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,789
; CURRENT FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sp-1 consensus
; OTHER INFORMATION: sequence
US-09-989-789-15

Query Match 73.4%; Score 45.5; DB 9; Length 100;
Best Local Similarity 42.9%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
   ||||| |||||
Db 19 CPEGKSFSSKSHLRAHQTH 39

RESULT 4
US-09-846-033B-223
; Sequence 223, Application US/09846033B
; Publication No. US20030044404A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolfe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; TITLE OF INVENTION: Finger Proteins
; FILE REFERENCE: 019496-005820US
; CURRENT APPLICATION NUMBER: US/09/846,033B
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 223
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sp-i consensus sequence with leader sequence
US-09-846-033B-223

Query Match 73.4%; Score 45.5; DB 11; Length 100;
Best Local Similarity 42.9%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
   ||||| |||||
Db 19 CPEGKSFSSKSHLRAHQTH 39

RESULT 5
US-09-990-186-15
; Sequence 15, Application US/09990186
; Publication No. US20030068675A1
; GENERAL INFORMATION:
; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.21 / S11-US3
; CURRENT APPLICATION NUMBER: US/09/990,186
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sp-1 consensus
; OTHER INFORMATION: sequence
US-09-990-186-15

Query Match 73.4%; Score 45.5; DB 11; Length 100;
Best Local Similarity 42.9%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
   ||||| |||||
Db 19 CPEGKSFSSKSHLRAHQTH 39

RESULT 6
US-09-989-994-15
; Sequence 15, Application US/09989994
; Publication No. US20030104528A1
; GENERAL INFORMATION:
```

; APPLICANT: LIU, Qiang
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE
; FILE REFERENCE: 8325-0011.20 / S11-US2
; CURRENT APPLICATION NUMBER: US/09/989,994
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ ID NOS: 4085
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sp-1 consensus
US-09-989-994-15

Query Match 73.4%; Score 45.5; DB 11; Length 100;
Best Local Similarity 42.9%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 19 CPEGKSFSSHLRAHQTH 39

RESULT 7
US-10-006-069A-223
; Sequence 223, Application US/10006069A
; Publication No. US20030021776A1
; GENERAL INFORMATION:
; APPLICANT: Rebar, Edward
; APPLICANT: Jamieson, Andrew
; APPLICANT: Liu, Qiang
; APPLICANT: Liu, Pei-Qi
; APPLICANT: Wolffe, Alan
; APPLICANT: Eisenberg, Stephen P.
; APPLICANT: Jarvis, Eric
; APPLICANT: Sangamo Biosciences, Inc.
; TITLE OF INVENTION: Regulation of Angiogenesis With Zinc
; FILE REFERENCE: 019496-005830US
; CURRENT APPLICATION NUMBER: US/10/006,069A
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 09/733,604
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 09/736,083
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 09/846,033
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 223
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sp-1 consensus sequence with leader sequence
US-10-006-069A-223

Query Match 73.4%; Score 45.5; DB 15; Length 100;
Best Local Similarity 42.9%; Pred. No. 4;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 19 CPEGKSFSSHLRAHQTH 39

RESULT 8
US-09-765-555-38
; Sequence 38, Application US/09765555
; Publication No. US20030037355A1

; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZFPm1
US-09-765-555-38

Query Match 73.4%; Score 45.5; DB 11; Length 184;
Best Local Similarity 42.9%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 154 CPEGKSFSGAGHLASHQTH 174

RESULT 9
US-09-765-555-39
; Sequence 39, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ZFPm2
US-09-765-555-39

Query Match 73.4%; Score 45.5; DB 11; Length 184;
Best Local Similarity 42.9%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 126 CPEGKSFSGAGHLASHQTH 146

RESULT 10
US-09-765-555-75
; Sequence 75, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40

; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Zinc finger protein zFPml
US-09-765-555-75

Query Match 73.4%; Score 45.5; DB 11; Length 184;
Best Local Similarity 42.9%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
Qy 1 CPEC-----XHQTH 10
| | | | |
Db 154 CPEGKSFQAGHLASHORTH 174

RESULT 11
US-09-908-1538-32
; Sequence 32, Application US/09/9081538
; Patent No. US20020168714A1
; GENERAL INFORMATION:
; APPLICANT: Bardas, Carlos F.
; APPLICANT: Beerli, Roger
; APPLICANT: Schopfer, Ulrich
; TITLE OF INVENTION: REGULATION OF GENE EXPRESSION USING
; TITLE OF INVENTION: SINGLE-CHAIN, MONOMERIC, LIGAND DEPENDENT POLYPEPTIDE
; TITLE OF INVENTION: SWITCHES
; FILE REFERENCE: TSRI 725.1
; CURRENT APPLICATION NUMBER: US/09/908,153B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 09/619,063
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Synthesized
; NAME/KEY: VARIANT
; LOCATION: 1, 186
; OTHER INFORMATION: Xaa - Any Amino Acid
US-09-908-1538-32

Query Match 73.4%; Score 45.5; DB 10; Length 186;
Best Local Similarity 42.9%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
Qy 1 CPEC-----XHQTH 10
| | | | |
Db 127 CPEGKSFQAGHLASHORTH 147

RESULT 12
US-10-198-677-77
; Sequence 77, Application US/10198677
; Publication No. US20030119023A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: KLOG, Aaron
; APPLICANT: MOORE, Michael
; TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY

; TITLE OF INVENTION: FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
; TITLE OF INVENTION: MOLECULES
; FILE REFERENCE: 8325-2011 / G11-US1
; CURRENT APPLICATION NUMBER: US/10/198,677
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 77
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus structure
US-10-198-677-77

Query Match 71.8%; Score 44.5; DB 15; Length 25;
Best Local Similarity 42.9%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
Qy 1 CPEC-----XHQTH 10
| | | | |
Db 4 CPEGKSFQKSDLVKHQTH 24

RESULT 13
US-09-851-271A-8
; Sequence 8, Application US/09851271A
; Patent No. US20020064824A1
; GENERAL INFORMATION:
; APPLICANT: Gendag Limited
; TITLE OF INVENTION: Screening System
; FILE REFERENCE: 674538-2003
; CURRENT APPLICATION NUMBER: US/09/851,271A
; CURRENT FILING DATE: 2001-05-08
; PRIOR APPLICATION NUMBER: PCT/GB99/03730
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: GB9824544.2
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: ZN_FING
; LOCATION: (1)..(26)
; OTHER INFORMATION: zinc finger consensus structure
US-09-851-271A-8

Query Match 71.8%; Score 44.5; DB 9; Length 26;
Best Local Similarity 42.9%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
Qy 1 CPEC-----XHQTH 10
| | | | |
Db 4 CPEGKSFQKSDLVKHQTH 24

RESULT 14
US-09-995-973-8
; Sequence 8, Application US/09959573
; Publication No. US20030024006A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: GENE SWITCHES
; FILE REFERENCE: 8325-2003 / G7-US1
; CURRENT APPLICATION NUMBER: US/09/995,973
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8

; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: structure
US-09-995-973-8

Query Match 71.8%; Score 44.5; DB 11; Length 26;
Best Local Similarity 42.9%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|
|
|
|
Db 4 CPEGKSFQKSDLVXHQTH 24

RESULT 15
US-09-996-484-8
; Sequence 8, Application US/09996484
; Publication No. US20030092010A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: MOLECULAR SWITCHES
; FILE REFERENCE: 8325-2004 / G8-US1
; CURRENT APPLICATION NUMBER: US/09/996,484
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: structure
US-09-996-484-8

Query Match 71.8%; Score 44.5; DB 11; Length 26;
Best Local Similarity 42.9%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|
|
|
|
Db 4 CPEGKSFQKSDLVXHQTH 24

Search completed: October 1, 2003, 10:50:51
Job time : 26.6667 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	45.5	73.4	98	4	US-09-229-007A-10	Sequence 10, Appl
2	44.5	71.8	26	2	US-08-491-527A-15	Sequence 15, Appl
3	44.5	71.8	26	2	US-08-620-151-49	Sequence 49, Appl
4	44.5	71.8	462	3	US-08-486-095-117	Sequence 117, App
5	44.5	71.8	462	3	US-08-360-107A-127	Sequence 127, App
6	44.5	71.8	462	3	US-08-484-223B-117	Sequence 117, App
7	44.5	71.8	462	3	US-08-919-597-117	Sequence 117, App
8	44.5	71.8	462	3	US-08-475-669A-117	Sequence 117, App
9	44.5	71.8	462	3	US-08-485-513A-117	Sequence 117, App
10	44.5	71.8	462	3	US-08-471-913A-117	Sequence 117, App
11	44.5	71.8	462	3	US-08-485-264A-117	Sequence 117, App
12	44.5	71.8	462	4	US-08-474-349A-117	Sequence 117, App
13	44.5	71.8	462	4	US-08-470-896-117	Sequence 117, App
14	44.5	71.8	462	4	US-08-485-546A-117	Sequence 117, App
15	44	71.0	279	1	US-08-300-903A-9	Sequence 9, Appl
16	44	71.0	279	4	US-08-988-197-9	Sequence 9, Appl
17	41	66.1	700	4	US-09-107-532A-5094	Sequence 5094, Ap
18	40	64.5	117	4	US-09-252-991A-17347	Sequence 17347, A
19	40	64.5	129	4	US-09-328-352-5196	Sequence 5196, Ap
20	39.5	63.7	26	2	US-08-620-151-27	Sequence 27, Appl
21	39.5	63.7	26	2	US-08-620-151-85	Sequence 85, Appl
22	39.5	63.7	675	1	US-08-317-522A-9	Sequence 9, Appl
23	39.5	63.7	675	1	US-08-439-818A-9	Sequence 9, Appl
24	39.5	63.7	675	2	US-08-751-965-9	Sequence 9, Appl
25	39.5	63.7	675	2	US-08-738-975-9	Sequence 9, Appl
26	39.5	63.7	675	2	US-08-728-626-9	Sequence 9, Appl
27	39.5	63.7	675	3	US-08-808-599A-9	Sequence 9, Appl

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/491,527A
FILING DATE: 16-JUNE-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/260,199
FILING DATE: 15-JUN-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/245,507
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 7900-0008.30
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: peptide CPI
US-08-491-527A-15

Query Match 71.8%; Score 44.5; DB 2; Length 26;
Best Local Similarity 42.9%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 4 CPEGKFSOKSGLVXHQTH 24
|||||

RESULT 3
US-08-620-151-49
Sequence 49, Application US/08620151
Patent No. 5928955
GENERAL INFORMATION:
APPLICANT: Imperiali, Barbara
APPLICANT: Walkup, Grant K.
TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR
DIVALENT ZINC
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOPER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/620,151
FILING DATE: 22-MAR-1996

CLASSIFICATION: 422
ATTORNEY/AGENT INFORMATION:
NAME: Shannon, Karen L.
REGISTRATION NUMBER: 36,675
REFERENCE/DOCKET NUMBER: 8597/6
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-620-151-49

Query Match 71.8%; Score 44.5; DB 2; Length 26;
Best Local Similarity 42.9%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 4 CPEGKGFQNSDLVXHQTH 24
|||||

RESULT 4
US-08-486-099-117
Sequence 117, Application US/08486099
Patent No. 6013263
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
B VIRUS TRANSMISSION
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
B VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,099
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-031
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 117:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-486-099-117

Query Match 71.8%; Score 44.5; DB 3; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 268 CPEGKSFROSTHLLHQTH 288

RESULT 5

US-08-360-107A-127
; Sequence 127, Application US/08360107A
; Patent No. 6017536
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,107A
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 127:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-360-107A-127

Query Match 71.8%; Score 44.5; DB 3; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

Qy 1 CPEC-----XHQTH 10
|||||
Db 268 CPEGKSFROSTHLLHQTH 288

RESULT 6

US-08-484-223B-117
; Sequence 117, Application US/08484223B
; Patent No. 6020459

; GENERAL INFORMATION:

; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-223B-117

Query Match 71.8%; Score 44.5; DB 3; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
|||||
Db 268 CPEGKSFROSTHLLHQTH 288

RESULT 7

US-08-919-597-117
; Sequence 117, Application US/08919597
; Patent No. 6054265
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

;
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,597
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470,896
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
; US-08-919-597-117

Query Match 71.8%; Score 44.5; DB 3; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHORTH 10
|||||
Db 268 CPEGKSFQSTHLLHQRTH 288
|||||

RESULT 8
US-08-475-668A-117
; Sequence 117, Application US/08475668A
; Patent No. 6060065
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,668A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742

;
; REFERENCE/DOCKET NUMBER: 7872-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
; US-08-475-668A-117

Query Match 71.8%; Score 44.5; DB 3; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHORTH 10
|||||
Db 268 CPEGKSFQSTHLLHQRTH 288
|||||

RESULT 9
US-08-485-551A-117
; Sequence 117, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,551A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: protein
; MOLECULE TYPE: protein
; US-08-485-551A-117

Query Match 71.8%; Score 44.5; DB 3; Length 462;

```
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
    ||||
    268 CPEGKSFQSTHLLHQTH 288

Db

RESULT 10
US-08-471-913A-117
; Sequence 117, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Pettaway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/471,913A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-471-913A-117

Query Match 71.8%; Score 44.5; DB 3; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
    ||||
    268 CPEGKSFQSTHLLHQTH 288

Db

RESULT 12
US-08-474-349A-117
; Sequence 117, Application US/08474349A
; Patent No. 6333595
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Pettaway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
; TITLE OF INVENTION: VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 517
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA

Query Match 71.8%; Score 44.5; DB 3; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
    ||||
    268 CPEGKSFQSTHLLHQTH 288

Db

RESULT 11
US-08-485-264A-117
; Sequence 117, Application US/08485264A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
```

```
;
;
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474.349A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-474-349A-117

Query Match 71.8%; Score 44.5; DB 4; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
    ||||
    ||||
Db 268 CPECCKSFQSTHLLHQTH 288

RESULT 13
US-08-470-896-117
; Sequence 117, Application US/08470896
; Patent No. 6479055
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470.896
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
```

```
;
;
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-470-896-117

Query Match 71.8%; Score 44.5; DB 4; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

QY 1 CPEC-----XHQTH 10
    ||||
    ||||
Db 268 CPECCKSFQSTHLLHQTH 288

RESULT 14
US-08-485-546A-117
; Sequence 117, Application US/08485546A
; Patent No. 6518013
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.546A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-546A-117

Query Match 71.8%; Score 44.5; DB 4; Length 462;
Best Local Similarity 42.9%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
```

```

QY      1 CPEC-----XHORT 10
Db      268 CPECCKSPRQSTHILHORTH 288

RESULT 15
US-08-300-903A-9
: Sequence 9, Application US/08300903A
: Patent No. 5591630
: GENERAL INFORMATION:
: APPLICANT: Anderson, Dirk M
: APPLICANT: Giri, Judith G
: TITLE OF INVENTION: Interleukin-15 Receptors
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Apple Operating System 7.1
: SOFTWARE: Microsoft Word for Apple, Version 5.1a
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/300,903A
: FILING DATE: 06-SEPTEMBER-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: USSN 08/236,919
: FILING DATE: 06-MAY-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Perkins, Patricia Anne
: REGISTRATION NUMBER: 34,695
: REFERENCE/DOCKET NUMBER: 2822-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-587-0430
: TELEFAX: 206-233-0644
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 279 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-300-903A-9

Query Match      71.0%; Score 44; DB 1; Length 279;
Best Local Similarity 66.7%; Pred No. 4,6;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 CPECXHQRT 9
Db      61 CPQCHHRT 69

```

Search completed: October 1, 2003, 10:49:09
Job time : 22.8333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:04:56 ; Search time 51.1111 Seconds
(without alignments)
24.844 Million cell updates/sec

Title: US-10-057-890A-4
Perfect score: 24
Sequence: 1 XXCXXXXC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

```
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	75.0	9	AA38591	Human lambda light
2	18	75.0	9	AAW58491	Human KOL antibody
3	18	75.0	14	AAW97990	Human peptide #126
4	18	75.0	14	AAW97991	Human peptide #126
5	18	75.0	15	AAW05980	Somatic immunoglob
6	18	75.0	15	ABJ26565	Guanylin analogue
7	18	75.0	18	ABP46759	Human BlyS binding
8	18	75.0	21	AAU02670	CDR region of anti
9	18	75.0	23	AAW07019	New mu-conotoxin p

10	18	75.0	23	22	AAE12409	Albumin fusion pro
11	18	75.0	24	20	AAV42462	Conotoxin peptide
12	18	75.0	24	21	AAW09054	Hepatitis GB virus
13	18	75.0	24	21	AAW92022	Toxin peptide seq
14	18	75.0	24	23	ABG99488	Conus sp conotoxin
15	18	75.0	24	23	ABG99801	Conus sp conotoxin
16	18	75.0	31	23	ABG99513	Conus sp conotoxin
17	18	75.0	31	23	ABG99793	Conus sp conotoxin
18	18	75.0	32	22	ABG59764	Human liver peptid
19	18	75.0	32	22	ABB44382	Peptide #11888 enco
20	18	75.0	32	22	ABB27241	Protein #9240 enco
21	18	75.0	32	22	AAW65457	Human brain expres
22	18	75.0	32	22	AAW78126	Human bone marrow
23	18	75.0	32	22	AAW22005	Peptide #8439 enco
24	18	75.0	32	22	AAW38362	Peptide #12399 enco
25	18	75.0	32	22	ABG48528	Human liver peptid
26	18	75.0	32	22	ABG49663	Human liver peptid
27	18	75.0	32	22	ABB28505	Peptide #1156 enco
28	18	75.0	32	22	ABB29653	Peptide #2304 enco
29	18	75.0	32	22	ABB33690	Peptide #1196 enco
30	18	75.0	32	22	ABB34829	Peptide #2335 enco
31	18	75.0	32	22	ABB19140	Peptide #2335 enco
32	18	75.0	32	22	ABB20249	Protein #1139 enco
33	18	75.0	32	22	AAW54458	Human brain expres
34	18	75.0	32	22	AAW53637	Human brain expres
35	18	75.0	32	22	AAW68663	Human bone marrow
36	18	75.0	32	22	AAW68017	Human bone marrow
37	18	75.0	32	22	AAW14725	Peptide #1159 enco
38	18	75.0	32	22	AAW15836	Peptide #2270 enco
39	18	75.0	32	22	AAW27148	Peptide #1185 enco
40	18	75.0	32	22	AAW28346	Peptide #2383 enco
41	18	75.0	32	22	AAW03448	Peptide #1130 enco
42	18	75.0	32	22	AAW03574	Peptide #2256 enco
43	18	75.0	32	22	ABG36520	Human peptide enco
44	18	75.0	32	23	ABG37546	Human peptide enco
45	18	75.0	34	23	AAU87022	Adenovirus 5 fiber

ALIGNMENTS

RESULT 1

AA38591
ID AAR38591 standard: peptide; 9 AA.

XX AAR38591;
AC AAR38591;

XX 25-MAR-2003 (updated)
DT 28-OCT-1993 (first entry)

XX Human lambda light chain subgroup 1 (hL1).

XX Antibody: variable domain; light; L; heavy; H; consensus;
KW affinity; antigen; immunogenicity; humanisation; framework.

OS Homo sapiens.

PN WO9311794-Al.

XX 24-JUN-1993.

XX 14-DEC-1992; 92WO-US10906.

XX 13-DEC-1991; 91US-0808464.

XX (XOMA) XOMA CORP.

XX Fishwild DM, Kohn FR, Little RG, Studnicka GM;

XX WPI; 1993-213827/26.

XX Antibodies prepn. used for treatment of auto-immune diseases - by
PT replacement of critical residues to reduce immunogenicity but

PT retain binding affinity, etc.

XX Claim 2; Page 92; 160pp; English.

XX The consensus amino acid sequences for the subgroups of light chains (HK1 - AAR38590, HK3 - NGK, HK2 - GST, HL1 - AAR38591, HL2 - AAR38592, HL3 - AAR38593, HL6 - AAR38594, HK4 - AAR38595, HL4 - AAR38596, and HL5 - AAR38597) and heavy chains (HH3 - AAR38598, HH1 - AAR38599, and HH2 - AAR38600) of human variable domains may be used to prepare, for example, a modified mouse antibody variable domain that retains the affinity of the natural domain for antigen while exhibiting reduced immunogenicity in humans.

XX Unlike other methods of humanisation, which advocate the replacement of entire antibody framework regions with those of human antibodies, this method involves only the introduction of human residues into those positions not critical for antigen binding. CC This ensures that the binding properties of the modified antibody are not diminished.

XX (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 9 AA;

Query Match 75.0%; Score 18; DB 14; Length 9;

Best Local Similarity 33.3%; Pred. No. 6.1e+05;

Matches 2; Conservative 0; Mismatches 4; Indels 0; Caps 0;

QY 3 CXXXXC 8

DB 4 CSSASC 9

RESULT 2

AAW58491

ID AAW58491 standard; protein; 9 AA.

XX AC AAW58491;

XX DT 18-AUG-1998 (first entry)

XX DE Human KOL antibody heavy chain variable domain fragment.

XX KW Humanised; human; mouse; CD5; anti-CD5 antibody; immunoglobulin; depletion; cytotoxic; immunoconjugate; fusion protein; psoriasis; autoimmune disease; rheumatoid arthritis; type I diabetes.

XX OS Homo sapiens.

XX PN US5770196-A.

XX PD 23-JUN-1998.

XX PF 07-JUN-1995; 95US-0472789.

XX PR 23-JUN-1993; 93US-0082842.

XX PR 13-DEC-1991; 91US-0808464.

XX PR 14-DEC-1992; 92WO-US10906.

XX PR 07-JUN-1995; 95US-0472788.

XX PA (XOMA) XOMA CORP.

XX PI Studnicka GM;

XX WP1; 1998-376744/32.

XX Depletion of CD5-positive cells in vivo - using anti-CD5 antibodies with humanised variable regions

XX Disclosure; Column 51; 77pp; English.

XX A method has been developed of depleting CD5+ cells in an animal. The method comprises administering a cytotoxic protein containing a modified immunoglobulin (Ig) variable domain, where the protein is an anti-CD5 Ig molecule or an immunoconjugate or fusion protein containing an anti-CD5

CC Ig molecule, and where the modified Ig variable domain comprises at least one of (a) a modified light chain variable region (see AAW58478 or AAW58480), and (b) a modified heavy chain variable region (see AAW58479 or AAW58481), where AAW58478 and AAW58479 are humanised forms of the H65 light and heavy chain variable domains with low risk amino acid substitutions [i.e. low risk of reducing antigen-binding specificity.] CC chain variable domains with moderate risk amino acid substitutions and are present in humanised H65 antibody he3 (ATCC HB 11206). The method is useful for treating autoimmune diseases, especially systemic lupus erythematosus, rheumatoid arthritis, psoriasis or type I diabetes. The present sequence represents the human KOL antibody heavy chain variable domain fragment.

XX Sequence 9 AA;

Query Match 75.0%; Score 18; DB 19; Length 9;

Best Local Similarity 33.3%; Pred. No. 6.1e+05;

Matches 2; Conservative 0; Mismatches 4; Indels 0; Caps 0;

QY 3 CXXXXC 8

DB 4 CSSASC 9

RESULT 3

AAW97990

ID AAW97990 standard; Peptide; 14 AA.

XX AC AAW97990;

XX DT 24-JAN-2002 (first entry)

XX DE Human peptide #1265 encoded by a SNP oligonucleotide.

XX KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic; neuroprotective; antimicrobial; gene therapy; vaccine; cancer; amyloid protein; angiotensin; apoptosis related protein; cadherin; cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor; complement related protein; cytochrome; kinase; cytokine; interferon; interleukin; G-protein coupled receptor; thioesterase; inflammation; multifactorial disease; autoimmune disease; infection; nervous system disease.

XX OS Homo sapiens.

XX PN WO200147944-A2.

XX PD 05-JUL-2001.

XX PF 28-DEC-2000; 2000WO-US35498.

XX PR 28-DEC-1999; 99US-0173419.

XX PR 27-DEC-2000; 2000US-0173419.

XX PA (CURA-) CURAGEN CORP.

XX PI Shinkets RA, Leach M;

XX WP1; 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases, oncogenes and histones, useful for diagnosing and treating, e.g. cancer, autoimmune diseases and infections

XX Disclosure; Page 3945; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659) encoding polymorphic variants of proteins related to amylases, amyloid proteins, angiotensin, apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes, histones, kinases, colony stimulating factors, complement related proteins, cytochromes, kinesins, cytokines, interferons, interleukins, G-protein coupled receptors and thioesterases.

CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.
 XX Sequence 14 AA;
 SQ Query Match 75.0%; Score 18; DB 22; Length 14;
 Best Local Similarity 33.3%; Pred. No. 1.8e+03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 CXXXXC 8
 DB 8 CTAATC 13
 RESULT 4
 ID AAM97991 standard; Peptide; 14 AA.
 XX
 AC AAM97991;
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE Human peptide #1266 encoded by a SNP oligonucleotide.
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
 KW amyloid protein; angioptietin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-US35498.
 XX
 PR 28-DEC-1999; 99US-0173419.
 PR 27-DEC-2000; 2000US-0173419.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI; 2001-465210/50.
 XX
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymorases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections.
 XX
 PS Disclosure; Page 3945; 4143pp; English.
 XX
 CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angioptietin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, G-protein coupled receptors, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with

CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.
 XX Sequence 14 AA;
 SQ Query Match 75.0%; Score 18; DB 22; Length 14;
 Best Local Similarity 33.3%; Pred. No. 1.8e+03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 CXXXXC 8
 DB 8 CTAATC 13
 RESULT 5
 ID AAR05980 standard; protein; 15 AA.
 XX
 AC AAR05980;
 XX
 DT 16-NOV-1990 (first entry)
 XX
 DE Somatic immunoglobulin DH fragment KOL.
 XX
 KW Immunoglobulin DH; cancer; lymphoid tumour; leukaemia; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO9006998-A.
 XX
 PD 28-JUN-1990.
 XX
 PF 19-DEC-1989; 89WO-JP01275.
 XX
 PR 19-DEC-1988; 88JP-0319809.
 XX
 PA (MITK) MITSUI TOATSU CHEM INC.
 PA (EDUC-) EDUCATIONAL FOUND FUJITA.
 XX
 PI KUROSAWA Y, ICHIHARA Y, AWAYA K, ISHIZUKA Y;
 XX
 DR WPI; 1990-224526/29.
 XX
 PT DNA fragments related to human immunoglobulin genes - in which
 PT are markers for diagnosis of lymphoid tumours by detection of B-
 PT and T-type tumour cells.
 XX
 PS Disclosure; Page ?; 91pp; Japanese.
 XX
 CC DH immunoglobulin encoding sequences may be extracted as fragments
 CC and used in diagnosis of lymphoid tumours and leukaemias, detecting
 CC marker sequences from the VhDhJh gene cluster.
 XX
 SQ Sequence 15 AA;
 Query Match 75.0%; Score 18; DB 11; Length 15;
 Best Local Similarity 33.3%; Pred. No. 1.9e+03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 CXXXXC 8
 DB 7 CSSASC 12
 RESULT 6
 ID ABJ26565 standard; Peptide; 15 AA.
 XX

AC ABJ26565;
 XX 10-APR-2003 (first entry)
 DT Guanylin analogue #13.
 DE
 XX Auditory; cytostatic; hypertensive; vasotropic; diuretic; natriuretic;
 KW cyclic; binding inhibitor; guanylin; guanylin cyclase C;
 KW Meniere's disease; fluctuating hearing level; roaring tinnitus;
 KW episodic vertigo; colorectal cancer; colonic carcinoma; diuresis;
 KW GC-C receptor; hypotension; erectile dysfunction; inner ear disorder.
 XX
 XX Unidentified.
 OS
 XX WO200279235-A2.
 PN 10-OCT-2002.
 XX
 XX 29-MAR-2002; 2002WO-1B02466.
 PF
 XX 30-MAR-2001; 2001DK-0000534.
 PR
 XX (ZEAL-) ZEALAND PHARMA AS.
 PA
 XX Larsen BD, Neve S, Qvortrup K, Petersen JS, Meier E;
 PI
 XX WPI: 2003-221246/21.
 DR
 XX New cyclic polypeptides useful for preparation of a medicament for
 PT treatment of at least one symptom of e.g. Menieres disease -
 XX
 XX Disclosure; Page 28; 96pp; English.
 PS
 XX The invention relates to novel cyclic polypeptides, which are capable of
 CC inhibiting binding of guanylin or a compound, which comprises guanylin
 CC like activity to guanylin cyclase C. The novel cyclic polypeptides are
 CC useful for the preparation of a medicament useful for treatment of at
 CC least one symptom of Meniere's disease e.g. fluctuating hearing levels,
 CC sensation of fullness in the ear, roaring tinnitus or episodic vertigo.
 CC The treatment is ameliorating, curative or prophylactic in individuals
 CC (preferably human). The novel cyclic polypeptides are also useful for the
 CC preparation of a medicament for the treatment of colorectal cancer
 CC (preferably colonic carcinoma). The novel cyclic polypeptides are also
 CC useful for facilitating diuresis in a mammal, for raising the antibody
 CC capable of binding specifically to the GC-C receptor or a fraction of its
 CC peptide sequence, for preventing or treating hypotension or erectile
 CC dysfunction and for the treatment of inner ear disorders. This sequence
 CC represents a guanylin analogue peptide relating to the invention.
 XX
 XX Sequence 15 AA:
 SQ
 Query Match 75.08; Score 18; DB 24; Length 15;
 Best Local Similarity 33.3%; Pred. No. 1.9e+03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 CXXXXC 8
 Db 7 CAAAAAC 12
 RESULT 7
 ABP46759
 ID ABP46759 standard; peptide; 18 AA.
 XX
 XX ABP46759;
 AC
 XX 19-AUG-2002 (first entry)
 DT
 XX Human Blys binding scFv VH CDR3 SEQ ID 2770.
 DE
 XX Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 OS

KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 XX Homo sapiens.
 OS
 XX WO200202641-A1.
 PN 10-JAN-2002.
 XX
 XX 15-JUN-2001; 2001WO-US19110.
 PF
 XX 16-JUN-2000; 2000US-212210P.
 PR
 XX 17-OCT-2000; 2000US-240816P.
 PR
 XX 21-MAR-2001; 2001US-276248P.
 PR
 XX 26-MAR-2001; 2001US-277379P.
 PR
 XX 25-MAY-2001; 2001US-293499P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 PI WPI: 2002-114799/15.
 DR
 XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for
 PT the diagnosis and treatment of cancers and immune disorders -
 XX
 XX Claim 2; Page 3048; 3148pp; English.
 PS
 XX This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID)) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method
 CC of the invention.
 XX
 XX Sequence 18 AA:
 SQ
 Query Match 75.08; Score 18; DB 23; Length 18;
 Best Local Similarity 33.3%; Pred. No. 1.9e+03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 CXXXXC 8
 Db 6 CSSTSC 11
 RESULT 8
 AAU02670
 ID AAU02670 standard; peptide; 21 AA.
 XX
 XX AAU02670;
 AC
 XX 29-AUG-2001 (first entry)
 DT
 XX CDR region of anti-adipocyte antibody heavy chain, FAT 31.
 DE
 XX Antibody; adipocyte; heavy chain; light chain; obesity; fat;
 KW heart disease; complementarity determining region; CDR.
 KW
 XX Homo sapiens.
 OS

XX WO200127279-A1.
 PN XX
 PD 19-APR-2001.
 XX
 PF 11-OCT-2000; 2000WO-GB03900.
 XX
 PR 12-OCT-1999; 99US-0158812.
 XX
 PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 PI Edwards BM, Main SH, Vaughan TJ;
 DR WPI; 2001-282031/29.
 DR N-PSDB; AAS03443.
 XX
 PT Panel of specific binding members of antibody molecules which bind to
 PT whole adipocytes is used in the treatment of obesity and obesity
 PT related diseases -
 XX
 PS Example 7; Page 77; 182pp; English.
 XX
 CC AAD02501-AAU02635, and AAU02641-AAU02748 represent the amino acid
 CC sequences of anti-adipocyte monoclonal antibody heavy chain, light
 CC chain, and heavy chain complementarity determining regions (CDR) of the
 CC invention. The antibodies can be used in the treatment of obesity and
 CC obesity related diseases. The antibodies can be used to deliver drugs or
 CC pro-drugs directly to the fat mass of an obese patient or the antibody
 CC can be used as a therapeutic itself. Antibodies binding specifically to
 CC adipocytes can be used to activate the immune system to destroy the cells
 CC by complement mediated lysis. The antibodies may be labeled with a
 CC detectable label such as radiolabel, fluorescent or chemical group and
 CC used in methods of diagnosis in human subjects e.g. to determine the
 CC presence of adipocyte antigen on the surface of an adipocyte to detect or
 CC determine the presence or level of adipocytes in a cell or tissue sample.
 CC The antibodies can be used as an alternative means of treatment for obese
 CC patients other than undergoing surgery to remove excess fat. Antibodies
 CC for different types of fat deposits can also be produced e.g. intra-
 CC abdominal fat associated with heart disease.
 XX
 XX Sequence 21 AA;

Query Match 75.0%; Score 18; DB 22; Length 21;
 Best Local Similarity 33.3%; Pred. No. 2e+03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 DB 9 CTSASC 14

RESULT 9
 AAR70719
 ID AAR70719 standard; peptide: 23 AA.

XX AAR70719;
 AC AAR70719;
 XX
 DT 25-MAR-2003 (updated)
 DT 25-SEP-1995 (first entry)
 XX
 DE New mu-conotoxin peptide which targets a specific receptor.
 XX
 KW conotoxin; disulphide bond; inhibitor; synaptic transmission;
 KW neuromuscular junction; class mu; skeletal muscle sodium ion channel.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 2
 FT Modified-site 3
 FT Modified-site 4
 FT Modified-site 14

FT /label= 4Hyp
 XX WO9501436-A1.
 PN XX
 PD 12-JAN-1995.
 XX
 PF 27-JUN-1994; 94WO-US07194.
 XX
 PR 29-JUN-1993; 93US-0084848.
 XX
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PA (UTAH) UNIV UTAH RES FOUND.
 XX
 PI Olivera BM, Rivier JEF, Cruz LJ, Abogadie F, Hopkins CE, Dykert J;
 PI Torres JL;
 XX
 DR WPI; 1995-061000/08.
 XX
 PT Conotoxin peptide(s) containing one or more cyclising di:sulphide
 PT bonds - inhibit synaptic transmissions at neuromuscular
 PT junctions, useful in binding assays and as pesticides
 XX
 PS Claim 18; Page 51; 56pp; English.
 XX
 CC This conotoxin peptide appears to be a member of the known class of
 CC mu-conotoxins. This class of conotoxins target and block the skeletal
 CC muscle sodium ion channels. The conotoxin peptides are useful
 CC as pesticides, and many of them or closely related analogues are
 CC targeted to specific insects or other pests.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 XX Sequence 23 AA;

Query Match 75.0%; Score 18; DB 16; Length 23;
 Best Local Similarity 33.3%; Pred. No. 2e+03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 DB 16 CSSASC 21

RESULT 10
 AAE12409
 ID AAE12409 standard; peptide: 23 AA.

XX AAE12409;
 XX
 DT 18-DEC-2001 (first entry)
 XX

DE Albumin fusion protein related human peptide.

XX
 KW Human; albumin; HA; immune system disorder; transplant rejection;
 KW blood related disorder; myocardial infarction; glomerulonephritis;
 KW hyperproliferative disorder; childhood acute myeloid leukaemia;
 KW renal cell carcinoma; cardiovascular disorder; vulnery; melanoma;
 KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukaemic;
 KW neurological disease; Alzheimer's disease; endocrine disorder; measles;
 KW pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;
 KW infectious disease; gastrointestinal disorder; wound healing; nontropic;
 KW irritable bowel syndrome; HIV; human immunodeficiency virus infection;
 KW cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiac;
 KW antiarthritic; antirheumatic; renal disorder; antimicrobial.

XX Homo sapiens.
 OS
 XX
 PN WO200179480-A1.
 XX
 PD 25-OCT-2001.
 XX
 PF 12-APR-2001; 2001WO-US11991.
 XX

XX (COGN-) COGNETIX INC.
 PA (UTAH) UNIV UTAH RES FOUND.
 XX Olivera BM, McIntosh JM, Watkins M, Garrett JE, Cruz LJ;
 PI Grilley M, Walker CS, Shetty R, Jones RM, Schoenfeld RM;
 XX WPI: 2002-706921/76.
 DR
 XX New cone snail conotoxin peptides, useful as a pain reliever for
 PT alleviating pain in an individual suffering from pain or who is about
 PT to be subjected to a pain-causing event, or for treating voltage-gated
 PT ion channel disorders -
 XX
 PS Claim 1; Page 179; 305pp; English.
 XX This invention describes novel conotoxin peptides from the cone snail,
 CC genus Conus which have analgesic activity and can act as a voltage-gated
 CC ion channel modulator or a ligand-gated ion channel modulator. The
 CC conotoxin peptide is useful as a pain-relieving agent for alleviating
 CC pain in an individual who is either exhibiting pain or is about to be
 CC subjected to a pain-causing event. The conotoxin peptide is also useful
 CC for treating or preventing disorders associated with voltage-gated ion
 CC channel disorders, ligand-gated ion channel disorders or receptor
 CC disorders. The radiolabeled conotoxin peptide is also useful for
 CC characterising a new site on these receptors or channels, and for
 CC screening and identifying novel small molecules that interact with the
 CC above-mentioned channels or receptors, which are monoamine transporters.
 CC ABG99360-ABG99853 represent the conotoxin protein and peptides described
 CC in the disclosure of the invention.
 XX
 SQ Sequence 24 AA;
 Query Match 75.0%; Score 18; DB 23; Length 24;
 Best Local Similarity 33.3%; Pred. No. 2.1e-03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Y 3 CXXXXC 8
 | |
 D 16 CSSASC 21
 RESULT 15
 ABG99801
) ABG99801 standard; Peptide; 24 AA.
 ABG99801;
 17-JAN-2003 (first entry)
 Conus sp conotoxin-associated peptide SEQ ID 586.
 Conotoxin; cone snail; analgesic; voltage-gated ion channel modulator;
 ligand-gated ion channel modulator; pain-relief.
 Conus consors.
 WO200264740-A2.
 22-AUG-2002.
 11-FEB-2002; 2002WO-US03887.
 09-FEB-2001; 2001US-267408P.
 (COGN-) COGNETIX INC.
 (UTAH) UNIV UTAH RES FOUND.
 Olivera BM, McIntosh JM, Watkins M, Garrett JE, Cruz LJ;
 Grilley M, Walker CS, Shetty R, Jones RM, Schoenfeld RM;
 WPI: 2002-706921/76.

PT New cone snail conotoxin peptides, useful as a pain reliever for
 PT alleviating pain in an individual suffering from pain or who is about
 PT to be subjected to a pain-causing event, or for treating voltage-gated
 PT ion channel disorders -
 XX
 PS Claim 1; Page 295; 305pp; English.
 XX This invention describes novel conotoxin peptides from the cone snail,
 CC genus Conus which have analgesic activity and can act as a voltage-gated
 CC ion channel modulator or a ligand-gated ion channel modulator. The
 CC conotoxin peptide is useful as a pain-relieving agent for alleviating
 CC pain in an individual who is either exhibiting pain or is about to be
 CC subjected to a pain-causing event. The conotoxin peptide is also useful
 CC for treating or preventing disorders associated with voltage-gated ion
 CC channel disorders, ligand-gated ion channel disorders or receptor
 CC disorders. The radiolabeled conotoxin peptide is also useful for
 CC characterising a new site on these receptors or channels, and for
 CC screening and identifying novel small molecules that interact with the
 CC above-mentioned channels or receptors, which are monoamine transporters.
 CC ABG99360-ABG99853 represent the conotoxin protein and peptides described
 CC in the disclosure of the invention.
 XX
 SQ Sequence 24 AA;
 Query Match 75.0%; Score 18; DB 23; Length 24;
 Best Local Similarity 33.3%; Pred. No. 2.1e-03;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 CXXXXC 8
 | |
 Db 16 CSSASC 21
 Search completed: October 1, 2003, 10:41:49
 Job time : 53.1111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

rotein - protein search, using sw model

on: October 1, 2003, 10:34:26 ; Search time 17.556 Seconds
(without alignments)
43.824 Million cell updates/sec

le: US-10-057-890A-4

fect score: 24

quence: 1 XXCXXXC 8

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 283308 seqs, 9616882 residues

tal number of hits satisfying chosen parameters: 283308

imum DB seq length: 0

imum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_76:*

2: PIR1:*

3: PIR2:*

4: PIR3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	18	75.0	23	153401	monocyte chemotact
2	18	75.0	24	A58999	psi-conotoxin PIIE
3	18	75.0	40	JT0515	Ig heavy chain V-I
4	18	75.0	86	256788	Ig heavy chain V r
5	18	75.0	88	S05527	capsicein - Phytol
6	18	75.0	98	S06671	cinnamomin - Phytol
7	18	75.0	99	S42360	capsicein - Phytol
8	18	75.0	113	B40278	neocarzinostatin -
9	18	75.0	113	T10136	hypothetical prote
10	18	75.0	118	S49913	cryptogain - Phytol
11	18	75.0	118	S49905	acidic elicitor Al
12	18	75.0	121	PH1661	Ig heavy chain V r
13	18	75.0	123	S30320	Ig heavy chain V r
14	18	75.0	123	S49906	HAE (highly acidic
15	18	75.0	123	S49907	HAE (highly acidic
16	18	75.0	126	GIHUKL	Ig heavy chain V-I
17	18	75.0	127	S24689	Ig heavy chain V r
18	18	75.0	128	S26788	Ig heavy chain V r
19	18	75.0	129	S36260	Ig heavy chain V r
20	18	75.0	134	D58335	Ly6 homolog RK3 pr
21	18	75.0	143	YNSMCG	actinoxanthin prec
22	18	75.0	147	T1NSMCC	neocarzinostatin p
23	18	75.0	183	S21495	hypothetical prote
24	18	75.0	204	T44257	thiamine-phosphate
25	18	75.0	210	T40465	related to high cy
26	18	75.0	230	T49555	probable mitosis a
27	18	75.0	237	G75316	probable lipase -
28	18	75.0	289	C96610	hypothetical prote
29	18	75.0	295	G91042	ethanolamine ammon

30	18	75.0	295	2	B85687	ethanolamine ammon
31	18	75.0	311	2	H82136	methyltetrahydr
32	18	75.0	316	1	S27981	homoserine kinase
33	18	75.0	321	2	T46352	hypothetical prote
34	18	75.0	326	1	J44891	glycoprotein PF7 p
35	18	75.0	334	2	JC4870	pepsin A (EC 3.4.23
36	18	75.0	337	2	A84913	hypothetical prote
37	18	75.0	352	2	A71447	probable WRK1-type
38	18	75.0	365	2	S68466	cathepsin E (EC 3.
39	18	75.0	374	2	T30429	late expression fa
40	18	75.0	387	2	E38302	pepsin (EC 3.4.23.
41	18	75.0	446	2	T50786	nucleoid DNA-bind
42	18	75.0	459	2	A25928	cellulase (EC 3.2.
43	18	75.0	463	2	A48375	endoglucanase I -
44	18	75.0	464	2	JC7143	rhodocoxin reducta
45	18	75.0	465	2	AD3579	

ALIGNMENTS

RESULT 1

I53401

monocyte chemotactic protein - human (fragment)

C.Species: Homo sapiens (man)

C.Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000

C.Accession: I53401

R.Steenbergen, E.J.; Verhagen, O.J.; van Leeuwen, E.F.; Behrendt, H.; Merle, P.A.; We

Eur. J. Immunol. 24, 900-908, 1994

A.Title: B precursor acute lymphoblastic leukemia third complementarity-determining r

fetal life.

A.Reference number: I53401; MUID:94200227; PMID:8149961

A.Accession: I53401

A.Status: Preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-23 <RES>

A.Cross-references: GB:S69742; NID:9546303; PIDN:AAD14040.1; PID:94261740

C.Genetics:

A.Gene: IGH-VDJ

Query Match 75.0%; Score 18; DB 2; Length 23;

Best Local Similarity 33.3%; Pred. No. 5.1e-02;

Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXC 8

DB 8 CSSTSC 13

RESULT 2

A58999

psi-conotoxin PIIE - cone shell (Conus purpurascens)

C.Species: Conus purpurascens (purple cone)

C.Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999

C.Accession: A58999

R.Shon, K.J.; Grilley, M.; Jacobsen, R.; Cartier, G.E.; Hopkins, C.; Gray, W.R.; Wat

Biochemistry 36, 9581-9587, 1997

A.Title: A noncompetitive peptide inhibitor of the nicotinic acetylcholine receptor

A.Reference number: A58999; MUID:97383165; PMID:9236004

A.Accession: A58999

A.Status: Preliminary

A.Molecule type: protein

A.Residues: 1-24 <SHO>

C.Comment: This conotoxin is a noncompetitive inhibitor of the nicotinic acetylcholine

C.Keywords: amidated carboxyl end; hydroxyproline; neurotoxin; venom

F.2,3,14/Modified site: 4-hydroxyproline (Pro) #status experimental

F.4-16,5-21,10-22/Disulfide bonds: #status experimental

F.24/Modified site: amidated carboxyl end (Arg) #status experimental

Query Match 75.0%; Score 18; DB 2; Length 24;

Best Local Similarity 33.3%; Pred. No. 5.1e-02;

Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 16 CSSASC 21

RESULT 3
JT0515
C:Species: Homo sapiens (man)
C:Date: 23-Oct-1992 #sequence_revision 23-Oct-1992 #text_change 30-May-1997
C:Accession: JT0515
R:Anker, R.; Conley, M.E.; Pollok, B.A.
J. Exp. Med. 169, 2109-2119, 1989
A:Title: Clonal diversity in the B cell repertoire of patients with X-linked agammaglobulinemia
A:Reference number: JT0511; MUID:89279157; PMID:2786547
A:Accession: JT0515
A:Molecule type: mRNA
A:Residues: 1-40 <ANK>
Note: The sequence shown here is one of eight productive V-D-J mu chain rearrangements
Superfamily: immunoglobulin V region; immunoglobulin homology
Keywords: heterotetramer; immunoglobulin
1-24/Domain: V region <VPE>
25-37/Domain: D region <DRE>
38-40/Domain: J region <JRE>

Query Match 75.0%; Score 18; DB 2; Length 40;
Best Local Similarity 33.3%; Pred. No. 5.9e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 29 CSSTSC 34

RESULT 4
788
heavy chain V region - human
Species: Homo sapiens (man)
Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Jun-2000
Accession: S26788
R:Ortali, F.; Newton, J.A.; Wang, J.Y.; Schroeder Jr., H.W.
J. Immunol. 22, 241-245, 1992
Title: The human cord blood antibody repertoire. Frequent usage of the V(H)7 gene family
Reference number: S26786; MUID:92111632; PMID:1730251
Accession: S26788
Status: preliminary
Molecule type: mRNA
Residues: 1-86 <MOR>
Cross-references: EMBL:X61017; NID:932788; PIDN:CAA43351.1; PID:g1335123
Superfamily: immunoglobulin V region; immunoglobulin homology
Keywords: heterotetramer; immunoglobulin

Query Match 75.0%; Score 18; DB 2; Length 86;
Best Local Similarity 33.3%; Pred. No. 7.2e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 63 CTSSTC 68

RESULT 5
7
zein - phytophthora capsici
Species: Phytophthora capsici
Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 18-Aug-2000
Accession: S05527
R:Bonnet, P.; Huet, J.C.; Sallantin, M.; Beauvais-Cante, F.; Bruneteau, M.; B
Biochem. 183, 555-563, 1989
Title: Structure and activity of proteins from pathogenic fungi Phytophthora eliciting
Reference number: S05526; MUID:89377822; PMID:2776750
Accession: S05527
Molecule type: protein

A:Residues: 1-98 <RIC>
C:Comment: This protein elicits leaf necrosis and acquired resistance in tobacco.
C:Superfamily: cryptogein X24

Query Match 75.0%; Score 18; DB 2; Length 98;
Best Local Similarity 33.3%; Pred. No. 7.4e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 51 CASTAC 56

RESULT 6
S06671
cinnamomin - Phytophthora cinnamomi
Species: Phytophthora cinnamomi
Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Aug-2000
Accession: S06671
R:Huet, J.C.; Pernollet, J.C.
FEBS Lett. 257, 302-306, 1989
Title: Amino acid sequence of cinnamomin, a new member of the elicitin family, and
Reference number: S06671; MUID:90060341; PMID:2583377
Accession: S06671
Molecule type: protein
Residues: 1-98 <HUE>
Superfamily: cryptogein X24

Query Match 75.0%; Score 18; DB 2; Length 98;
Best Local Similarity 33.3%; Pred. No. 7.4e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 51 CASTAC 56

RESULT 7
S42360
capsicein - phytophthora capsici
Species: Phytophthora capsici
Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 18-Aug-2000
Accession: S42360
R:Bouaziz, S.; van Heijenoort, C.; Guittet, E.; Huet, J.C.; Pernollet, J.C.
Eur. J. Biochem. 220, 427-438, 1994
Title: Resonance assignment, cysteine-pairing elucidation and secondary-structure d
Reference number: S42360; MUID:94170791; PMID:8125100
Accession: S42360
Status: preliminary
Molecule type: protein
Residues: 1-99 <BOU>
Superfamily: cryptogein X24

Query Match 75.0%; Score 18; DB 2; Length 99;
Best Local Similarity 33.3%; Pred. No. 7.4e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 52 CASTAC 57

RESULT 8
B40278
neocarzinostatin - Streptomyces malayensis
Alternate names: mitomycin
Species: Streptomyces malayensis
Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
Accession: B40278
R:Gao, X.; Burkhardt, W.
Biochemistry 30, 7730-7739, 1991
Title: Two- and three-dimensional proton NMR studies of apo-neocarzinostatin.
Reference number: A40278; MUID:91329336; PMID:1831044

A:Accession: B40278
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-113 <GAO>
 C:Superfamily: macromycin
 C:Keywords: antibiotic; DNA binding; pigment binding
 F:37-47,88-93/disulfide bonds: #status predicted

Query Match 75.0%; Score 18; DB 1; Length 113;
 Best Local Similarity 33.3%; Pred. No. 7.7e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 Db 88 CTTAAC 93

RESULT 9

hypothetical protein mrh.1 - phage T4

C:Species: phage T4
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
 C:Accession: T10136; JH0139
 R:Prazier, M.W.; Mosig, G.
 Gene 88, 7-14, 1990
 A:Title: The bacteriophage T4 gene mrh whose product inhibits late T4 gene expression in
 A:Reference number: JH0136; MUID:90255970; PMID:1692800
 A:Accession: T10136
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-113 <FRA>
 A:Cross-references: EMBL:M30001; NID:g1532074; PID:g1532078
 A:Experimental source: strain G17
 C:Genetics:
 A:Gene: mrh.1
 A:Map position: 14.692 - 15.031 kb
 C:Superfamily: phage T4 hypothetical protein mrh.1

Query Match 75.0%; Score 18; DB 2; Length 113;
 Best Local Similarity 33.3%; Pred. No. 7.7e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 Db 107 CSSTSC 112

RESULT 10

S49913

cryptogaein - Phytophthora cryptogae

C:Species: Phytophthora cryptogae
 C:Date: 05-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 23-Mar-2001
 C:Accession: S49913; S05576; S49908
 R:Pambieres, F.; Marais, A.; Le Berre, J.; Penot, I.; Fournier, D.; Ricci, P.
 submitted to the EMBL Data Library, June 1994

A:Description: Coding sequence of the gene encoding cryptogaein.

A:Reference number: S49906
 A:Accession: S49913
 A:Molecule type: DNA

A:Residues: 1-118 <PAN>

A:Cross-references: EMBL:Z34459; NID:g599952; PIDN:CAA84224.1; PID:g599953
 R:Ricci, P.; Bonnet, P.; Huet, J.C.; Sallantin, M.; Beauvais-Cante, F.; Bruneteau, M.; B
 Eur. J. Biochem. 183, 555-563, 1989

A:Title: Structure and activity of proteins from pathogenic fungi Phytophthora eliciting

A:Reference number: S05526; MUID:89377822; PMID:2776750

A:Accession: S05526

A:Molecule type: protein

A:Residues: 21-118 <RIC>

C:Comment: This protein elicits leaf necrosis and acquired resistance in tobacco.

C:Genetics:

A:Gene: X24

C:Superfamily: cryptogaein X24

Query Match 75.0%; Score 18; DB 2; Length 118;
 Best Local Similarity 33.3%; Pred. No. 7.8e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 Db 71 CASTAC 76

RESULT 11

S49905

acidic elicitin A1 - Phytophthora cryptogae

C:Species: Phytophthora cryptogae
 C:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 18-Aug-2000
 C:Accession: S49905
 R:Pambieres, F.; Marais, A.; Le Berre, J.; Penot, I.; Fournier, D.; Ricci, P.
 submitted to the EMBL Data Library, June 1994

A:Description: Coding sequence of the gene encoding acidic isoform of cryptogaein A1.

A:Reference number: S49905

A:Accession: S49905

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-118 <PAN>

A:Cross-references: EMBL:Z34462; NID:g599946; PIDN:CAA84227.1; PID:g599947

C:Genetics:

A:Gene: B14

C:Superfamily: cryptogaein X24

Query Match 75.0%; Score 18; DB 2; Length 118;
 Best Local Similarity 33.3%; Pred. No. 7.8e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 Db 71 CASTAC 76

RESULT 12

PH1661

Ig heavy chain V region (clone SJI) - human (fragment)

C:Species: Homo sapiens (man)
 C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 16-Aug-1996
 C:Accession: PH1661
 R:Hillson, J.L.; Karr, N.S.; Opplinger, I.R.; Mannik, M.; Sasso, E.H.
 J. Exp. Med. 178, 331-336, 1993

A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to stap

A:Reference number: PH1642; MUID:93301610; PMID:8315388

A:Accession: PH1661

A:Molecule type: mRNA

A:Residues: 1-121 <HIL>

A:Experimental source: B cell

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:7-90/Domain: immunoglobulin homology <IMM>

Query Match 75.0%; Score 18; DB 2; Length 121;
 Best Local Similarity 33.3%; Pred. No. 7.8e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 Db 94 CSSSSC 99

RESULT 13

S30530

Ig heavy chain V region - human

C:Species: Homo sapiens (man)
 C:Date: 03-Mar-1994 #sequence_revision 10-Nov-1995 #text_change 16-Aug-1996
 C:Accession: S30530

R:Mariette, X.

submitted to the EMBL Data Library, October 1992

A:Reference number: S30520

Search completed: October 1, 2003, 10:47:48
Job time : 19.5556 secs

A:Accession: S30510
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-123 <MAR>
A:Cross-references: EMBL:218316
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 75.0%; Score 18; DB 2; Length 123;
Best Local Similarity 33.3%; Pred. No. 7.9e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 CXXXXC 8
DB 102 CSSTSC 107

RESULT 14
S49906
HAE (highly acidic elicitin) 20 - Phytophthora cryptogea
C:Species: Phytophthora cryptogea
C>Date: 05-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 23-Mar-2001
C:Accession: S49906
R:Panabieres, F.; Marais, A.; Le Berre, J.; Penot, I.; Fournier, D.; Ricci, P.
submitted to the EMBL Data Library, June 1994
A:Description: Coding sequence of the gene encoding hypothetical HAE (highly acidic elicitin)
A:Reference number: S49907
A:Accession: S49906
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <PAN>
A:Cross-references: EMBL:234460; NID:g599948; PIDN:CAA84225.1; PID:g599949
C:Genetics:
A:Gene: B20
C:Superfamily: cryptogein X24

Query Match 75.0%; Score 18; DB 2; Length 123;
Best Local Similarity 33.3%; Pred. No. 7.9e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 CXXXXC 8
DB 71 CASTAC 76

RESULT 15
S49907
HAE (highly acidic elicitin) 26 - Phytophthora cryptogea
C:Species: Phytophthora cryptogea
C>Date: 05-Mar-1995 #sequence_revision 14-Jul-1995 #text_change 23-Mar-2001
C:Accession: S49907
R:Panabieres, F.; Marais, A.; Le Berre, J.; Penot, I.; Fournier, D.; Ricci, P.
submitted to the EMBL Data Library, June 1994
A:Description: Coding sequence of the gene encoding hypothetical HAE (highly acidic elicitin)
A:Reference number: S49908
A:Accession: S49907
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <PAN>
A:Cross-references: EMBL:234461; NID:g599950; PIDN:CAA84226.1; PID:g599951
C:Genetics:
A:Gene: B26
C:Superfamily: cryptogein X24

Query Match 75.0%; Score 18; DB 2; Length 123;
Best Local Similarity 33.3%; Pred. No. 7.9e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Y 3 CXXXXC 8
b 71 CASTAC 76

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:33:32 : Search time 47.7778 Seconds
(without alignments)
43.209 Million cell updates/sec

Title: US-10-057-890A-4
Perfect score: 24
Sequence: 1 XXCXXXC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP_TREMBL23:*
- 2: sp_arChaea:*
- 3: sp_bacteria:*
- 4: sp_fungi:*
- 5: sp_human:*
- 6: sp_invertebrate:*
- 7: sp_mammal:*
- 8: sp_mhc:*
- 9: sp_organelle:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	75.0	70	2	Q05719 azospirillum
2	18	75.0	77	10	Q8L6X4 beta coroll
3	18	75.0	87	10	Q8HSC5 oryza sativ
4	18	75.0	92	10	Q9AT01 phytophthor
5	18	75.0	98	10	Q9S9F8 phytophthor
6	18	75.0	98	10	Q96731 phytophthor
7	18	75.0	98	10	Q96732 phytophthor
8	18	75.0	98	10	Q9S879 phytophthor
9	18	75.0	98	10	Q96734 phytophthor
10	18	75.0	98	10	Q96733 phytophthor
11	18	75.0	98	12	Q8VAB8 white spot
12	18	75.0	102	6	Q9BE83 macaca fasc
13	18	75.0	104	5	Q9U521 caenorhabdi
14	18	75.0	104	10	Q8L6X4 arabidopsis
15	18	75.0	104	11	Q8BQA8 mus musculu
16	18	75.0	111	11	Q8C606 mus musculu

17	18	75.0	112	4	Q9BYD5	Q9byd5 homo sapien
18	18	75.0	114	10	Q43225	Q43225 tulipa gesn
19	18	75.0	118	10	Q01905	Q01905 phytophthor
20	18	75.0	118	10	Q8L691	Q8L691 phytophthor
21	18	75.0	118	10	Q9T0W7	Q9T0W7 phytophthor
22	18	75.0	118	10	Q8L690	Q8L690 phytophthor
23	18	75.0	118	10	Q97360	Q97360 phytophthor
24	18	75.0	118	16	Q8RPE3	Q8rfe3 fusobacteri
25	18	75.0	122	10	Q97361	Q97361 phytophthor
26	18	75.0	123	12	Q11435	Q11435 duck adenov
27	18	75.0	123	12	Q8QTG9	Q8qtg9 white spot
28	18	75.0	124	5	Q8T5W5	Q8t5w5 caenorhabdi
29	18	75.0	124	10	Q97359	Q97359 phytophthor
30	18	75.0	126	5	Q8T5W4	Q8t5w4 caenorhabdi
31	18	75.0	127	5	Q8T5X2	Q8t5x2 caenorhabdi
32	18	75.0	128	4	Q96SH5	Q96sh5 homo sapien
33	18	75.0	128	5	Q8T5W8	Q8t5w8 caenorhabdi
34	18	75.0	128	10	Q8L6Q3	Q8l6q3 arabidopsis
35	18	75.0	129	5	Q8T5W7	Q8t5w7 caenorhabdi
36	18	75.0	133	15	Q91104	Q91104 human immun
37	18	75.0	134	4	Q9H8L9	Q9h8l9 homo sapien
38	18	75.0	134	11	Q63318	Q63318 rattus norv
39	18	75.0	146	10	Q8S348	Q8s348 capsicum an
40	18	75.0	152	5	Q8SW66	Q8sw66 encephalito
41	18	75.0	158	11	Q8BM99	Q8bm99 mus musculu
42	18	75.0	163	2	Q93NB8	Q93nb8 myxococcus
43	18	75.0	163	5	Q811H4	Q811h4 drosophila
44	18	75.0	167	4	Q96N15	Q96n15 homo sapien
45	18	75.0	172	8	Q9G6S6	Q9g6s6 sardinops m

ALIGNMENTS

RESULT 1

Q05719 PRELIMINARY; PRT; 70 AA.
ID Q05719 PRELIMINARY; PRT; 70 AA.
AC Q05719; 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE NifB (Fragment).
GN NIFB.
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Azospirillum.
OX NCBI_TaxID=192;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93122522; PubMed=1362170;
RA Liang Y., de Zamaroczy M., Arsene F., Paquelin A., Elmerich C.;
RT "Regulation of nitrogen fixation in Azospirillum brasilense Sp7:
RL Involvement of nifA, glbA and glbB gene products.";
RL FEMS Microbiol. Lett. 79:113-119(1992).
DR EMBL; S52383; AAB24913.1;
FT NON_TER 70 70
SQ SEQUENCE 70 AA; 7237 MW; AC87D0451D865C84 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 70;
Best Local Similarity 33.3%; Pred. No. 82;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8

Db 31 CASSSC 36

RESULT 2

Q8L6X4 PRELIMINARY; PRT; 77 AA.
ID Q8L6X4 PRELIMINARY; PRT; 77 AA.
AC Q8L6X4;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Reverse transcriptase (Fragment).
GN POL.
OS Beta corolliflora.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
OX NCBI_TaxID=19767;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Desel C.;
RT Fluorescence in situ hybridisation on chromosomes of nematode
RT resistant sugar beet lines.*;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ489204; CAD43266.1; -
DR InterPro: IPR003006; IG_MHC.
DR PROSITE: P500290; IG_MHC; 1.
FT NON_TER 1 1
FT NON_TER 77 77
SQ SEQUENCE 77 AA; 9172 MW; A34EAE0EBDAE2884 CRC64;
Query Match 75.0%; Score 18; DB 10; Length 77;
Best Local Similarity 33.3%; Pred. No. 82;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 CXXXXC 8
Db 72 CYSSTC 77
RESULT 3
ID Q8H5C5 PRELIMINARY; PRT; 87 AA.
AC Q8H5C5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE QJ1343_B12.33 protein.
GN QJ1343_B12.33.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
clone:QJ1343_B12.33";
RI Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003824; BAC19972.1; -
SQ SEQUENCE 87 AA; 9683 MW; CCA29D575723B791 CRC64;
Query Match 75.0%; Score 16; DB 10; Length 87;
Best Local Similarity 33.3%; Pred. No. 83;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 CXXXXC 8
Db 78 CSTSAC 83
RESULT 4
ID Q9AT01 PRELIMINARY; PRT; 92 AA.
AC Q9AT01;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE Sojein 1 protein (Fragment).
OS Phytophthora sojae.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora capsici.

OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4784;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1319;
RA Bailey A.N., Munoz C.I., Martinez P., Manriquez M.;
RT "Detection of additional gene products induced by the interaction
between Phytophthora capsici and its host, Capsicum annuum, and the
identification of an elicitor-encoding gene.*";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF361169; AAK32727.1; -
DR HSP; P15570; 1BEO.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS: PR00948; ELICITIN.
FT NON_TER 1 1
FT NON_TER 92 92
SQ SEQUENCE 92 AA; 9776 MW; E6E3FF429E889AD4 CRC64;
Query Match 75.0%; Score 18; DB 10; Length 92;
Best Local Similarity 33.3%; Pred. No. 84;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 CXXXXC 8
Db 46 CASTAC 51
RESULT 5
ID Q9S9F8 PRELIMINARY; PRT; 98 AA.
AC Q9S9F8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Elicitin 310.
OS Phytophthora parasitica (Potato buckeye rot agent).
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4792;
RN [1]
RP SEQUENCE.
RX MEDLINE=95226052; PubMed=7766061;
RA Mouton-Perronnet F., Bruneteau M., Denoroy L., Bouliteau P., Ricci P.,
RA Bonnet P., Michel G.;
RI Phytochemistry 38:41-44(1995).
DR HSP; P15570; 1BEO.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS: PR00948; ELICITIN.
SQ SEQUENCE 98 AA; 10356 MW; 6BE83C539B6DAC5E CRC64;
Query Match 75.0%; Score 18; DB 10; Length 98;
Best Local Similarity 33.3%; Pred. No. 84;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 CXXXXC 8
Db 51 CASTAC 56
RESULT 6
ID Q96731 PRELIMINARY; PRT; 98 AA.
AC Q96731;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Sojein 1 protein (Fragment).
OS Phytophthora sojae.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.

```

OX NCBI_TaxID=67593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=race 1;
RA Becker J., Nagel S., Tenhaken R.;
RT "Cloning, expression and characterization of protein elicitors from
RT the soybean pathogenic fungus Phytophthora sojae.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ007858; CAA07710.1; -.
DR HSSP: P15570; LBEO.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS: PR00948; ELICITIN.
FT NON_TER 1
SQ SEQUENCE 98 AA; 10184 MW; E6FBB6CADD0D0A71 CRC64;

Query Match 75.0%; Score 18; DB 10; Length 98;
Best Local Similarity 33.3%; Pred.No. 84;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
Db 51 CASTAC 56

RESULT 7
O96732 ID O96732 PRELIMINARY; PRT; 98 AA.
AC O96732;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Sojein 2 protein (Fragment).
OS Phytophthora sojae.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=67593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=race 1;
RA Becker J., Nagel S., Tenhaken R.;
RT "Cloning, expression and characterization of protein elicitors from
RT the soybean pathogenic fungus Phytophthora sojae.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ007859; CAA07711.1; -.
DR HSSP: P15570; LBEO.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS: PR00948; ELICITIN.
FT NON_TER 1
SQ SEQUENCE 98 AA; 10123 MW; 892D5044AEC36248 CRC64;

Query Match 75.0%; Score 18; DB 10; Length 98;
Best Local Similarity 33.3%; Pred.No. 84;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
Db 51 CASTAC 56

RESULT 8
O9S879 ID O9S879 PRELIMINARY; PRT; 98 AA.
AC O9S879;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Alpha elicitin, INFESTIN.
OS Phytophthora infestans (Potato late blight fungus).
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4787;

```

```

RN [1]
RP SEQUENCE.
RX MEDLINE=94281678; PubMed=8012047;
RA Huet J.C., Salle-Tourne M., Pernollet J.C.;
RT "Amino acid sequence and toxicity of the alpha elicitin secreted with
RT ubiquitin by Phytophthora infestans.";
RL Mol. Plant Microbe Interact. 7:302-304(1994).
DR HSSP: P15570; LBEO.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS: PR00948; ELICITIN.
SQ SEQUENCE 98 AA; 10332 MW; 392F2E7EDCF7F3F CRC64;

Query Match 75.0%; Score 18; DB 10; Length 98;
Best Local Similarity 33.3%; Pred.No. 84;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
Db 51 CASTAC 56

RESULT 9
O96734 ID O96734 PRELIMINARY; PRT; 98 AA.
AC O96734;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Sojein 4 protein (Fragment).
OS Phytophthora sojae.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=67593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=race 1;
RA Becker J., Nagel S., Tenhaken R.;
RT "Cloning, expression and characterization of protein elicitors from
RT the soybean pathogenic fungus Phytophthora sojae.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ007861; CAA07713.1; -.
DR HSSP: P15570; LBEO.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS: PR00948; ELICITIN.
FT NON_TER 1
SQ SEQUENCE 98 AA; 10126 MW; 4A34B6CADFD0DA73 CRC64;

Query Match 75.0%; Score 18; DB 10; Length 98;
Best Local Similarity 33.3%; Pred.No. 84;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
Db 51 CASTAC 56

RESULT 10
O96733 ID O96733 PRELIMINARY; PRT; 98 AA.
AC O96733;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Sojein 3 protein (Fragment).
OS Phytophthora sojae.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=67593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=race 1;

```

RA Becker J., Nagel S., Tenhaken R.;
 RT "Cloning, expression and characterization of protein elicitors from
 the soybean pathogenic fungus *Phytophthora sojae*.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ007860; CAA07112.1; -;
 DR HSSP: P15570; 1BEO.

DR InterPro: IPRO02200; Elicitin.
 DR Pfam: PF00964; Elicitin.1.
 DR PRINTS: PRO0948; ELICITIN.
 FT NON_TER 1
 SQ SEQUENCE 98 AA; 10181 MW; 25E25044AC11B24A CRC64;

Query Match 75.0%; Score 18; DB 10; Length 98;
 Best Local Similarity 33.3%; Pred. No. 84;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 Db 51 CASTAC 56

RESULT 11

O8VAB8 ID Q8VAB8 PRELIMINARY; PRT; 98 AA.

AC Q8VAB8;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Wsv507.

OS White spot syndrome virus (WSSV).

OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.

OX NCBI_TaxID=92852;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=21548311; PubMed=11689662;

RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;

RT "Complete genome sequence of the shrimp white spot bacilliform

virus.";

RL J. Virol. 75:11811-11820(2001).

RN [2]

RP SEQUENCE FROM N.A.

RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF332093; AAL33508.1; -;

SQ SEQUENCE 98 AA; 11468 MW; D1E884B15ECDF34F CRC64;

Query Match 75.0%; Score 18; DB 12; Length 98;
 Best Local Similarity 33.3%; Pred. No. 84;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 Db 57 CSSTAC 62

RESULT 12

O9BE83 ID Q9BE83 PRELIMINARY; PRT; 102 AA.

AC Q9BE83;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE Hypothetical 10.7 kDa protein.

OS Macaca fascicularis (Crab eating macaque) (Cynomolpus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Macaca.

OX NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Cerebellum;

RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,

RA Suzuki Y., Sugano S., Hashimoto K.;

RT "Isolation of full-length cDNA clones from macaque brain cDNA
 libraries.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB056763; BAB39311.1; -;
 KW Hypothetical protein.

SQ SEQUENCE 102 AA; 10670 MW; B4E75207196269EC CRC64;

Query Match 75.0%; Score 18; DB 6; Length 102;
 Best Local Similarity 33.3%; Pred. No. 85;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 Db 70 CAAAC 75

RESULT 13

O9U5Z1 ID Q9U5Z1 PRELIMINARY; PRT; 104 AA.

AC Q9U5Z1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 13, Last annotation update)

DE Transposase (Fragment).

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RW7406; TRANSPOSON=Tc2;

RX MEDLINE=92189726; PubMed=1312336;

RA Ruvolo V., Hill J.E., Levitt A.;

RT "The Tc2 transposon of *Caenorhabditis elegans* has the structure of a

self-regulated element.";

RL DNA Cell Biol. 11:111-122(1992).

DR EMBL: X59156; CAB56745.1; -;

FT NON_TER 1

SQ SEQUENCE 104 AA; 11977 MW; 4FE81C463A79E529 CRC64;

Query Match 75.0%; Score 18; DB 5; Length 104;
 Best Local Similarity 33.3%; Pred. No. 85;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 Db 79 CSSSC 84

RESULT 14

O8LBK1 ID Q8LBK1 PRELIMINARY; PRT; 104 AA.

AC Q8LBK1;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical protein.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,

RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;

RT "Full-length messenger RNA sequences greatly improve genome

annotation.";

RL Genome Biol. 0:0-0(2002).

RN [2]

RP SEQUENCE FROM N.A.

RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,

RA Feldmann K.;

RT "Full-Length cDNA from Arabidopsis thaliana."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY087164; AAM64720.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 104 AA; 11357 MW; C86F5B04D1D1DCA CRC64;

Query Match 75.0%; Score 18; DB 10; Length 104;
 Best Local Similarity 33.3%; Pred. No. 85;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 CXXXXC 8
 | |
 Db 64 CSSTTC 69

RESULT 15

Q8BOA8 PRELIMINARY; PRT; 104 AA.
 AC Q8BOA8; 2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK051132; BAC34534.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 104 AA; 11872 MW; 03B4799C95B765C2 CRC64;

Query Match 75.0%; Score 18; DB 11; Length 104;
 Best Local Similarity 33.3%; Pred. No. 85;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 CXXXXC 8
 | |
 Db 67 CTTTAC 72

Search completed: October 1, 2003, 10:46:24
 Job time : 50.7778 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:06:46 ; Search time 9,3333 Seconds
(without alignments)
40.309 Million cell updates/sec

Title: US-10-057-890A-4
Perfect score: 24
Sequence: 1 XXCXXXC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	75.0	27	CX3E_CONPU	P56529 conus purpu
2	18	75.0	98	ELIA_PHYCP	P15571 phytophthor
3	18	75.0	98	ELIA_PHYDR	P35696 phytophthor
4	18	75.0	98	ELIA_PHYNE	P35698 phytophthor
5	18	75.0	98	ELIB_PHYCI	P15569 phytophthor
6	18	75.0	98	ELIB_PHYDR	P35697 phytophthor
7	18	75.0	98	ELIB_PHYNE	P35699 phytophthor
8	18	75.0	113	Y011_BPT4	P20706 bacterioph
9	18	75.0	118	ELI1_PHYCR	P41802 phytophthor
10	18	75.0	118	ELIB_PHYCR	P15570 phytophthor
11	18	75.0	118	ELI1_PHYPR	P41801 phytophthor
12	18	75.0	123	ELI2_PHYCR	P41803 phytophthor
13	18	75.0	123	ELI3_PHYCR	P41804 phytophthor
14	18	75.0	126	HV3K_HUMAN	P01772 homo sapien
15	18	75.0	143	CAGA_STRLG	O06110 streptomyc
16	18	75.0	147	NCZS_STRCZ	P01550 streptomyc
17	18	75.0	204	THIE_RH1ET	O34294 rhizobium e
18	18	75.0	210	MOB1_SCHPO	O94360 schizosacch
19	18	75.0	252	NUC1_CUNEE	P81203 cunningham
20	18	75.0	295	EUTC_ECO57	O8xeh3 escherichia
21	18	75.0	326	V509_ROT2	O03874 equine rota
22	18	75.0	334	LCM1_HUMAN	O9uic8 homo sapien
23	18	75.0	337	R223_ARATH	O22900 arabidopsis
24	18	75.0	387	PEP3_RABIT	P27822 oryctolagus
25	18	75.0	440	AROA_SYNEL	O8dly3 synechococc
26	18	75.0	459	GUN1_TRIRE	P07981 trichoderma
27	18	75.0	463	GUN1_TRILLO	O12714 trichoderma
28	18	75.0	469	APTE_DROME	P29673 drosophila
29	18	75.0	473	KCC4_HUMAN	O15566 homo sapien
30	18	75.0	481	HL_DROHY	P56674 drosophila
31	18	75.0	482	DUSA_HUMAN	O9y6w6 homo sapien
32	18	75.0	483	DUSA_MOUSE	O9ess0 mus musculu
33	18	75.0	485	NAB1_MESAU	O35589 mesocricetu

34	18	75.0	485	1	NAB1_RAT	O62722 rattus norv
35	18	75.0	486	1	NAB1_HUMAN	Q13506 homo sapien
36	18	75.0	486	1	NAB1_MOUSE	Q61122 mus musculu
37	18	75.0	492	1	TISD_HUMAN	P47974 homo sapien
38	18	75.0	493	1	NIFB_RHISN	Q53205 rhizobium s
39	18	75.0	502	1	NIFB_AZOVI	P11067 azotobacter
40	18	75.0	539	1	Y469_HUMAN	O9u1p4 homo sapien
41	18	75.0	557	1	C791_SORBI	Q43135 sorghum bic
42	18	75.0	623	1	ZPI_MOUSE	Q62005 mus musculu
43	18	75.0	638	1	GLE_CHLRE	P31178 chlamydomon
44	18	75.0	803	1	FSPO_XENLA	P35447 xenopus lae
45	18	75.0	807	1	FSPO_RAT	P35446 rattus norv

ALIGNMENTS

RESULT 1						
CX3E_CONPU						
ID	CX3E_CONPU	STANDARD;	PRT;	27	AA.	
AC	P56529;					
DT	15-JUL-1998 (Rel. 36, Created)					
DT	30-MAY-2000 (Rel. 39, Last sequence update)					
DT	28-FEB-2003 (Rel. 41, Last annotation update)					
DE	Psi-conotoxin PIIE precursor (Fragment).					
OS	Conus purpurascens (Purple cone).					
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;					
OC	Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;					
OC	Neogastropoda; Conoidea; Conidae; Conus.					
OX	NCBI_TaxID=41690;					
RN	[1]					
RP	SEQUENCE FROM N.A., SEQUENCE OF 3-26, AND MASS SPECTROMETRY.					
RC	TISSUE=Venom:					
RX	MEDLINE=97383165; PubMed=9236004;					
RA	Shon K.-J., Grille M., Jacobsen R.B., Cartier G.E., Hopkins C.,					
RA	Gray W.R., Watkins M., Hilliard D.R., Rivier J.E., Torres J.,					
RA	Yoshikami D., Olivera B.M.;					
RT	*A noncompetitive peptide inhibitor of the nicotinic acetylcholine					
RT	receptor from Conus purpurascens venom.;"					
RL	Biochemistry 36:9581-9587(1997).					
RN	[2]					
RP	STRUCTURE BY NMR OF 3-26					
RX	MEDLINE=98138433; PubMed=9477946;					
RA	Mitchell S.S., Shon K.-J., Foster M.P., Davis D.R., Olivera B.M.,					
RA	Ireland C.M.;					
RT	*Three-dimensional solution structure of conotoxin psi-PIIE, an					
RT	acetylcholine gated ion channel antagonist.;"					
RL	Biochemistry 37:1215-1220(1998).					
CC	-!- FUNCTION: ACTS ON POSTSYNAPTIC MEMBRANES, BINDS AS A					
CC	NONCOMPETITIVE INHIBITOR TO THE NICOTINIC ACETYLCHOLINE RECEPTORS					
CC	(NACHR) AND THUS INHIBITS THEM.					
CC	-!- SUBCELLULAR LOCATION: Secreted.					
CC	-!- TISSUE SPECIFICITY: Expressed by the venom duct.					
CC	-!- MASS SPECTROMETRY: MW=2716.0; METHOD=LSIMS; RANGE=3-26.					
CC	-!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS.					
DR	PDB; IAS; 14-OCT-98.					
KW	Postsynaptic neurotoxin; Neurotoxin; Toxin;					
KW	Acetylcholine receptor inhibitor; Amidation; Hydroxylation;					
KW	3D-structure.					
FT	NON_TER	1	2			
FT	PROPEP	<1	26			
FT	PEPTIDE	3	18			
FT	DISULFID	6	18			
FT	DISULFID	7	23			
FT	DISULFID	12	24			
FT	MOD_RES	4	4			
FT	MOD_RES	5	5			
FT	MOD_RES	16	16			
FT	MOD_RES	26	26			
FT	VARIANT	8	8			
FT	VARIANT	11	11			
FT	VARIANT	26	26			
FT	TURN	9	10			

FT TURN 16 17
 FT TURN 19 20
 SQ SEQUENCE 27 AA: 3046 MW: 33E3DC137C1A2A9D CRC64;
 Query Match 75.0%; Score 18; DB 1; Length 27;
 Best Local Similarity 33.3%; Pred. No. 1.4e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 DB 18 CSSASC 23

RESULT 2
 ELIA_PHYCP STANDARD; PRT; 98 AA.
 AC P15571;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Alpha-elicitin capsicin.
 OS Phytophthora capsici.
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Phytophthora.
 OX NCBI_TaxID=4784;
 RN [1]
 RP SEQUENCE.
 RC STRAIN-Isolate 147;
 RX MEDLINE=89377822; PubMed=2776750;
 RA Ricci P., Bonnet P., Huet J.-C., Sallantin M., Beauvais-Cante F.,
 RA Brunetau M., Billard V., Michel G., Pernollet J.-C.;
 RT "Structure and activity of proteins from pathogenic fungi
 RT Phytophthora eliciting necrosis and acquired resistance in tobacco.";
 RL Eur. J. Biochem. 183:555-563(1989).
 RN [2]
 RP STRUCTURE BY NMR.
 RX MEDLINE=94170791; PubMed=8125100;
 RA Bouaziz S., van Heijenoort C., Guittet E., Huet J.-C.,
 RA Pernollet J.-C.;
 RT "Resonance assignment, cysteine-pairing elucidation and secondary-
 RT structure determination of capsicin, an alpha-elicitin, by three-
 RT dimensional 1H NMR";
 RL Eur. J. Biochem. 220:427-438(1994).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=94304827; PubMed=8031752;
 RA Bouaziz S., van Heijenoort C., Huet J.-C., Pernollet J.-C.,
 RA Guittet E.;
 RT "1H and 15N resonance assignment and secondary structure of
 RT capsicin, an alpha-elicitin, determined by three-dimensional
 RT heteronuclear NMR";
 RL Biochemistry 33:8188-8197(1994).
 CC -1- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
 CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
 CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
 CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH
 CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
 DR PIR: S05327; S05527.
 DR PIR: S42360; S42360.
 DR HSSP: P15570; 1BEO.
 DR PIR: S42360; S42360.
 DR PIR: S42360; S42360.
 DR InterPro: IPR002200; Elicitin.
 DR Pfam: PF00964; Elicitin; 1.
 DR PRINTS: PR00948; ELICITIN.
 DR DISULFID 3 71
 FT DISULFID 27 56
 FT DISULFID 51 95
 SQ SEQUENCE 98 AA: 10166 MW: A3B75CE27CB177AD CRC64;

Query Match 75.0%; Score 18; DB 1; Length 98;
 Best Local Similarity 33.3%; Pred. No. 2.2e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 DB 51 CASTAC 56

RESULT 3
 ELIA_PHYDR STANDARD; PRT; 98 AA.
 AC P35696;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Alpha-elicitin DRE-alpha.
 OS Phytophthora drechsleri.
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Phytophthora.
 OX NCBI_TaxID=4794;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92344801; PubMed=1368359;
 RA Huet J.-C., Nespoulous C., Pernollet J.-C.;
 RA "Structures of elicitin isoforms secreted by Phytophthora
 RA drechsleri.";
 RL Phytochemistry 31:1471-1476(1992).
 CC -1- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
 CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
 CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
 CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH
 CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
 DR HSSP: P15570; 1BEO.
 DR InterPro: IPR002200; Elicitin.
 DR Pfam: PF00964; Elicitin; 1.
 DR PRINTS: PR00948; ELICITIN.
 DR DISULFID 3 71
 FT DISULFID 27 56
 FT DISULFID 51 95
 SQ SEQUENCE 98 AA: 10328 MW: 244457CADB86A55B CRC64;

Query Match 75.0%; Score 18; DB 1; Length 98;
 Best Local Similarity 33.3%; Pred. No. 2.2e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 DB 51 CSSTAC 56

RESULT 4
 ELIA_PHYME STANDARD; PRT; 98 AA.
 AC P35698;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Alpha-elicitin MGM-alpha.
 OS Phytophthora megasperma (Potato pink rot fungus).
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Phytophthora.
 OX NCBI_TaxID=4788;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93319720; PubMed=7763784;
 RA Huet J.-C., Pernollet J.-C.;
 RA "Sequences of acidic and basic elicitin isoforms secreted by
 RA Phytophthora megasperma megasperma.";
 RL Phytochemistry 33:797-805(1993).
 CC -1- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
 CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
 CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
 CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH

CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
 DR HSSP; P15570; 13EO.
 DR InterPro; IPR002200; Elicitin.
 DR Pfam; PF00964; Elicitin; 1.
 DR PRINTS; PR00948; ELICITIN.
 FT DISULFID 3 71 BY SIMILARITY.
 FT DISULFID 27 56 BY SIMILARITY.
 FT DISULFID 51 95 BY SIMILARITY.
 SQ SEQUENCE 98 AA; 10239 MW; 34FFF79FD16363E1 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 98;
 Best Local Similarity 33.3%; Pred. No. 2.2e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 Db 51 CASTAC 56

RESULT 5

ELIB_PHYCI STANDARD; PRT; 98 AA.
 AC P15569;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Beta-elicitin cinnamomin.
 OS Phytophthora cinnamomi.
 CC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 CC Phytophthora.
 OX NCBI_TaxID=4785;
 RN [1]

RP SEQUENCE.

RX MEDLINE-90060341; PubMed-2583277;
 RA Huet J.-C.; Pernollet J.-C.;
 RT "Amino acid sequence of cinnamomin, a new member of the elicitin family, and its comparison to cryptogein and capsicin.";
 RL FEBS Lett. 257:302-306(1989).
 CC -!- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.

CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
 DR PIR; S06671; S06671.
 DR PDB; 1LJP; 31-JUL-02.
 DR InterPro; IPR002200; Elicitin.
 DR Pfam; PF00964; Elicitin; 1.
 DR PRINTS; PR00948; ELICITIN.

KW 3D-structure.

FT DISULFID 3 71 BY SIMILARITY.
 FT DISULFID 27 56 BY SIMILARITY.
 FT DISULFID 51 95 BY SIMILARITY.
 SQ SEQUENCE 98 AA; 10294 MW; 600DA52057CB46B CRC64;

Query Match 75.0%; Score 18; DB 1; Length 98;
 Best Local Similarity 33.3%; Pred. No. 2.2e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 Db 51 CASTAC 56

RESULT 6

ELIB_PHYDR STANDARD; PRT; 98 AA.
 AC P35697;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Beta-elicitin DRE-beta.
 OS Phytophthora drechsleri.
 CC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 CC Phytophthora.
 OX NCBI_TaxID=4794;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-92344801; PubMed-1368359;
 RA Huet J.-C.; Nespoulous C.; Pernollet J.-C.;
 RT "Structures of elicitin isoforms secreted by Phytophthora drechsleri.";
 RL Phytochemistry 31:1471-1476(1992).
 CC -!- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
 DR HSSP; P15570; 13EO.
 DR InterPro; IPR002200; Elicitin.
 DR Pfam; PF00964; Elicitin; 1.
 DR PRINTS; PR00948; ELICITIN.
 FT DISULFID 3 71 BY SIMILARITY.
 FT DISULFID 27 56 BY SIMILARITY.
 FT DISULFID 51 95 BY SIMILARITY.
 FT VARIANT 2 2 T->A
 SQ SEQUENCE 98 AA; 10339 MW; EF6A8CEB2A7300C8 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 98;
 Best Local Similarity 33.3%; Pred. No. 2.2e+02;
 Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
 | |
 Db 51 CASTAC 56

RESULT 7

ELIB_PHYME STANDARD; PRT; 98 AA.
 AC P35699;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Beta-elicitin MCM-beta.
 OS Phytophthora megasperma (Potato pink rot fungus).
 CC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 CC Phytophthora.
 OX NCBI_TaxID=4788;
 RN [1]

RP SEQUENCE.

RX MEDLINE-93319720; PubMed-7763784;

RA Huet J.-C.; Pernollet J.-C.;
 RT "Sequences of acidic and basic elicitin isoforms secreted by Phytophthora megasperma megasperma.";
 RL Phytochemistry 33:797-805(1993).
 CC -!- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.

DR HSSP; P15570; 13EO.

DR InterPro; IPR002200; Elicitin.

DR Pfam; PF00964; Elicitin; 1.

DR PRINTS; PR00948; ELICITIN.

FT DISULFID 3 71 BY SIMILARITY.

FT DISULFID 27 56 BY SIMILARITY.

FT DISULFID 51 95 BY SIMILARITY.

SQ SEQUENCE 98 AA; 10379 MW; 5160D4043E58C0F CRC64;

Query Match 75.0%; Score 18; DB 1; Length 98;
Best Local Similarity 33.3%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 51 CASTAC 56

RESULT 8
ID Y011_BPT4 STANDARD; PRT; 113 AA.
AC P20706: Q96225;
DT 01-FEB-1991 (Rel. 17, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 12.6 kDa protein in mhr-soc intergenic region.
GN Y011 OR MRH.1
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90255970; PubMed=1692800;
RA Frazier M.W., Mosig G.;
RT "The bacteriophage T4 gene mhr whose product inhibits late T4 gene
expression in an Escherichia coli rhoH (sigma 32) mutant.";
RL Gene 88:7-14(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT "Bacteriophage T4 genome analysis.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M30001; AAB07795.1; --
DR EMBL: AF158101; AAD42613.1; --
DR PIR: T10136; T10136.
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 12621 MW; 5F289787FD1E2F54 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 113;
Best Local Similarity 33.3%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 107 CSSTSC 112

RESULT 9
ID ELIB_PHYCR STANDARD; PRT; 118 AA.
AC P41802;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Acidic elicitin A1 precursor.
GN B14.
OS Phytophthora cryptogea.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4786;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Isolate 52;
RX MEDLINE=96112753; PubMed=8664508;
RA Panabieres F., Marais A., Le Berre J.Y., Penot I., Fournier D.,
RA Ricci P.;
RT "Characterization of a gene cluster of Phytophthora cryptogea which
RT codes for elicitors, proteins inducing a hypersensitive-like response
RT in tobacco.";
RL Mol. Plant Microbe Interact. 8:996-1003(1995).
CC -!- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH
CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z34462; CAA84227.1; --
DR PIR: S49905; S49905.
DR HSSP: P15570; LBEO.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS: PR00948; ELICITIN.
KW Signal; Multigene family.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 118 ACIDIC ELICITIN A1.
FT DISULFID 23 91 BY SIMILARITY.
FT DISULFID 47 76 BY SIMILARITY.
FT DISULFID 71 115 BY SIMILARITY.
SQ SEQUENCE 118 AA; 12181 MW; 114C97E3FF8ED2F7 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 118;
Best Local Similarity 33.3%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 71 CASTAC 76

RESULT 10
ID ELIB_PHYCR STANDARD; PRT; 118 AA.
AC P15570;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Beta-elicitin cryptogean precursor (CRY).
OS Phytophthora cryptogea.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4786;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Isolate 52;
RA Panabieres F., Marais A., le Berre J., Penot I., Fournier D.,
RA Ricci P.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 21-118.
RC STRAIN=Isolate 52;
RX MEDLINE=89377822; PubMed=2776750;
RA Ricci P., Bonnet P., Huet J.-C., Sallantin M., Beauvais-Cante F.,
RA Brunetau M., Billard V., Michel G., Pernollet J.-C.;

```

RT "Structure and activity of proteins from pathogenic fungi
RT Phytophthora eliciting necrosis and acquired resistance in tobacco.";
RL Eur. J. Biochem. 183:555-563(1989).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=97148337; PubMed=8994969;
RA Bolssy G., de la Fortelle E., Kahn R., Huet J.-C., Bricogne G.,
RA Pernollet J.-C., Brunie S.;
RT "Crystal structure of a fungal elicitor secreted by Phytophthora
RT cryptogaea, a member of a novel class of plant necrotic proteins.";
RL Structure 4:1429-1439(1996).
RN [4]
RN STRUCTURE BY NMR.
RX MEDLINE=98046740; PubMed=9385630;
RA Fefeau S., Bouaziz S., Huet J.-C., Pernollet J.-C., Guittet E.;
RT "Three-dimensional solution structure of beta cryptogein, a beta
RT elicitin secreted by a phytopathogenic fungus Phytophthora
RT cryptogaea.";
RL Protein Sci. 6:2279-2284(1997).
CC -!- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH
CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z34459; CA84224.1; -
DR PIR; S49913; S49913.
DR PDB; 1BEO; 15-MAY-97.
DR PDB; 1BEG; 03-DEC-97.
DR PDB; 1BXM; 15-JUN-99.
DR PDB; 1LRI; 28-AUG-02.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS; PR00948; ELICITIN.
KW Signal; Multigene family; 3D-structure.
FT CHAIN 1 20 BETA-ELICITIN CRYPTOGEIN.
FT DISULFID 23 91
FT DISULFID 47 76
FT DISULFID 71 115
FT STRAND 22 22
FT STRAND 25 39
FT HELIX 40 40
FT TURN 40 40
FT HELIX 42 51
FT TURN 55 57
FT TURN 64 72
FT HELIX 74 86
FT STRAND 92 94
FT TURN 96 98
FT STRAND 101 102
FT HELIX 104 116
FT TURN 117 118
SQ SEQUENCE 118 AA; 12193 MW; 94DFD872F247D71 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 118;
Best Local Similarity 33.3%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 CXXXXC 8
DB 71 CASTAC 76

RT "Structure and activity of proteins from pathogenic fungi
RT Phytophthora eliciting necrosis and acquired resistance in tobacco.";
RL Eur. J. Biochem. 183:555-563(1989).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=97148337; PubMed=8994969;
RA Bolssy G., de la Fortelle E., Kahn R., Huet J.-C., Bricogne G.,
RA Pernollet J.-C., Brunie S.;
RT "Crystal structure of a fungal elicitor secreted by Phytophthora
RT cryptogaea, a member of a novel class of plant necrotic proteins.";
RL Structure 4:1429-1439(1996).
RN [4]
RN STRUCTURE BY NMR.
RX MEDLINE=98046740; PubMed=9385630;
RA Fefeau S., Bouaziz S., Huet J.-C., Pernollet J.-C., Guittet E.;
RT "Three-dimensional solution structure of beta cryptogein, a beta
RT elicitin secreted by a phytopathogenic fungus Phytophthora
RT cryptogaea.";
RL Protein Sci. 6:2279-2284(1997).
CC -!- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH
CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z34459; CA84224.1; -
DR PIR; S49913; S49913.
DR PDB; 1BEO; 15-MAY-97.
DR PDB; 1BEG; 03-DEC-97.
DR PDB; 1BXM; 15-JUN-99.
DR PDB; 1LRI; 28-AUG-02.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS; PR00948; ELICITIN.
KW Signal; Multigene family; 3D-structure.
FT CHAIN 1 20 BETA-ELICITIN CRYPTOGEIN.
FT DISULFID 23 91
FT DISULFID 47 76
FT DISULFID 71 115
FT STRAND 22 22
FT STRAND 25 39
FT HELIX 40 40
FT TURN 40 40
FT HELIX 42 51
FT TURN 55 57
FT TURN 64 72
FT HELIX 74 86
FT STRAND 92 94
FT TURN 96 98
FT STRAND 101 102
FT HELIX 104 116
FT TURN 117 118
SQ SEQUENCE 118 AA; 12193 MW; 94DFD872F247D71 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 118;
Best Local Similarity 33.3%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 CXXXXC 8
DB 71 CASTAC 76

```

```

RESULT 11
ELI2_PHYCPR STANDARD; PRT; 118 AA.
ID ELI2_PHYCPR
AC P41801;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Elicitin precursor.
GN PAR1.
OS Phytophthora parasitica (Potato buckeye rot agent).
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4792;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN=Isolate P1979;
RX MEDLINE=94100579; PubMed=8274771;
RA Kamoun S., Klucher K.M., Coffey M.D., Tyler B.M.;
RT "A gene encoding a host-specific elicitor protein of Phytophthora
RT parasitica.";
RL Mol. Plant Microbe Interact. 6:573-581(1993).
CC -!- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH
CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S67432; AAB29433.1; -
DR HSSP; P15570; 1BEO.
DR InterPro: IPR002200; Elicitin.
DR Pfam: PF00964; Elicitin; 1.
DR PRINTS; PR00948; ELICITIN.
KW Signal.
FT CHAIN 1 20 ELICITIN.
FT DISULFID 23 91 BY SIMILARITY.
FT DISULFID 47 76 BY SIMILARITY.
FT DISULFID 71 115 BY SIMILARITY.
SQ SEQUENCE 118 AA; 12336 MW; 780D86FE95100DD3 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 118;
Best Local Similarity 33.3%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 CXXXXC 8
DB 71 CASTAC 76

RESULT 12
ELI2_PHYCPR STANDARD; PRT; 123 AA.
ID ELI2_PHYCPR
AC P41803;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Highly acidic elicitin 20 precursor.
GN B20.
OS Phytophthora cryptogaea.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4786;
RN [1]

```

```
RP SEQUENCE FROM N.A.
RC STRAIN-Isolate 52;
RX MEDLINE-96112753; PubMed-8664508;
RA Panabieres F., Marais A., Le Berre J.Y., Penot I., Fournier D.,
RA Ricci P.;
RT *Characterization of a gene cluster of Phytophthora cryptogea which
RT codes for elicitors, proteins inducing a hypersensitive-like response
RL in tobacco.*;
RL Mol. Plant Microbe Interact. 8:996-1003(1995).
CC -1- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH
CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: Z34460; CAA84225.1; -.
CC PIR: S49906; S49906.
CC HSSP: P15570; IBE0.
CC InterPro: IPR002200; Elicitin.
CC Pfam: PF00964; Elicitin; 1.
CC PRINTS: PR00948; ELICITIN.
CC Signal: Multigene family.
CC SIGNAL 1 20 BY SIMILARITY.
CC CHAIN 21 123 HIGHLY ACIDIC ELICITIN 20.
CC DISULFID 23 91 BY SIMILARITY.
CC DISULFID 47 76 BY SIMILARITY.
CC DISULFID 71 115 BY SIMILARITY.
CC SEQUENCE 123 AA; 12695 MW; ECD8F55BEDAE2D CRC64;
Query Match 75.0%; Score 18; DB 1; Length 123;
Best Local Similarity 31.3%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 3 CXXXXC 8
DB 71 CASTAC 76
RESULT 13
ELI3_PHYCR
ID ELI3_PHYCR STANDARD; PRT; 123 AA.
AC P41804;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Highly acidic elicitin 26 precursor.
GN B26.
OS Phytophthora cryptogea.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OC NCBI_TaxID=4786;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Isolate 52;
RX MEDLINE-96112753; PubMed-8664508;
RA Panabieres F., Marais A., Le Berre J.Y., Penot I., Fournier D.,
RA Ricci P.;
RT *Characterization of a gene cluster of Phytophthora cryptogea which
RT codes for elicitors, proteins inducing a hypersensitive-like response
RL in tobacco.*;
RL Mol. Plant Microbe Interact. 8:996-1003(1995).
CC -1- FUNCTION: INDUCES LOCAL AND DISTAL DEFENSE RESPONSES (INCOMPATIBLE
CC HYPERSENSITIVE REACTION) IN PLANTS FROM THE SOLANACEAE AND
```

```
CC CRUCIFERAE FAMILIES. ELICIT LEAF NECROSIS AND CAUSE THE
CC ACCUMULATION OF PATHOGENESIS-RELATED PROTEINS. MIGHT INTERACT WITH
CC THE LIPIDIC MOLECULES OF THE PLASMA MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ELICITIN FAMILY.
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: Z34461; CAA84226.1; -.
CC PIR: S49907; S49907.
CC HSSP: P15570; IBE0.
CC InterPro: IPR002200; Elicitin.
CC Pfam: PF00964; Elicitin; 1.
CC PRINTS: PR00948; ELICITIN.
CC Signal: Multigene family.
CC SIGNAL 1 20 BY SIMILARITY.
CC CHAIN 21 123 HIGHLY ACIDIC ELICITIN 26.
CC DISULFID 23 91 BY SIMILARITY.
CC DISULFID 47 76 BY SIMILARITY.
CC DISULFID 71 115 BY SIMILARITY.
CC SEQUENCE 123 AA; 12572 MW; 00DFDEF2A78A59DB CRC64;
Query Match 75.0%; Score 18; DB 1; Length 123;
Best Local Similarity 33.3%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 3 CXXXXC 8
DB 71 CASTAC 76
RESULT 14
HV3K_HUMAN
ID HV3K_HUMAN STANDARD; PRT; 126 AA.
AC P01772;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V-III region KOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE-83289131; PubMed-6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT *Three-dimensional structure determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.*;
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE-81072295; PubMed-7441755;
RA Marquart M., Deisenhofer J., Huber R., Palm W.;
RT "Crystallographic refinement and atomic models of the intact
RT immunoglobulin molecule KOL and its antigen-binding fragment at 3.0 A
RT and 1.0-A resolution".
RL J. Mol. Biol. 141:369-391(1980).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR: A02055; G1HUKL.
DR PDB: 2FB4; 12-JUL-89.
DR PDB: 2IG2; 12-JUL-89.
DR GO: 0005576; C:extracellular; NAS.
DR GO: 0003823; F:antigen binding activity; NAS.
DR GO: 0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
```

DR InterPro: IPR003596; I9-V.
DR Pfam: PF00047; I9; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS00835; IG-LIKE; 1.
KW Immunoglobulin V region; 3D-structure; Pyrrolidone carboxylic acid.
FT DOMAIN 1 112
FT MOD_RES 1 1
FT DISULFID 22 96
FT DISULFID 105 110
FT STRAND 3 7
FT STRAND 11 12
FT TURN 14 15
FT STRAND 18 25
FT HELIX 29 31
FT STRAND 34 39
FT TURN 41 42
FT STRAND 45 51
FT TURN 53 54
FT STRAND 58 60
FT HELIX 62 64
FT STRAND 65 65
FT TURN 66 67
FT STRAND 68 73
FT TURN 74 77
FT STRAND 78 83
FT HELIX 88 90
FT STRAND 92 99
FT STRAND 106 106
FT TURN 107 108
FT STRAND 109 109
FT STRAND 113 116
FT STRAND 120 124
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 13718 MW; E4D71B52B16F8776 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 126;
Best Local Similarity 33.3%; Pred. No. 2.4e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
Db 105 CSSASC 110

RESULT 15
CAGA_STRGL STANDARD; PRT; 143 AA.
AC Q06110;
DI 01-FEB-1995 (Rel. 31, Created)
DI 01-FEB-1995 (Rel. 31, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Antitumor antibiotic C-1027 apoprotein precursor (C-1027-AG).
GN CAGA.
OS Streptomyces globisporus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI-TaxID:1908;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-1027;
RX MEDLINE-93113067; PubMed=1369059;
RA Sakata N., Ikono S., Hori M., Hamada M., Otani T.;
RT "Cloning and nucleotide sequencing of the antitumor antibiotic C-1027 apoprotein gene."
RT apoprotein gene."
RL Biosci. Biotechnol. Biochem. 56:1592-1595(1992).
RN [2]
RP SEQUENCE OF 34-143.
RX MEDLINE-9129278; PubMed=1368592;
RA Otani T., Yasuhara T., Minami Y., Shimazu T., Zhang R., Xie M.Y.;
RT "Purification and primary structure of C-1027-AG, a selective antagonist of antitumor antibiotic C-1027, from Streptomyces globisporus."
RT Agric. Biol. Chem. 55:407-417(1991).

RN [3]
RP CHARACTERIZATION.
RX MEDLINE-93277832; PubMed=8504075;
RA Sugiura Y., Matsumoto T.;
RT Some characteristics of DNA strand scission by macromolecular
RT antitumor antibiotic C-1027 containing a novel enediyne
RT chromophore.";
RL Biochemistry 32:5548-5553(1993).
CC -!- FUNCTION: BINDS NON-COVALENTLY TO A CHROMOPHORE WHICH IS THE
CC CYTOTOXIC AND MUTAGENIC COMPONENT OF THE ANTIBIOTIC. THE
CC CHROMOPHORE BINDS TO DNA AS A WEAK INTERCALATOR AND CAUSES SINGLE-
CC AND DOUBLE-STRAND BREAKS.
CC -!- SIMILARITY: BELONGS TO THE NEOCARZINOSTATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10827; BAA01609.1; -;
DR PIR: JCI315; YMSMCG.
DR PDB: 1H2K; 23-MAY-01.
DR PDB: 1H2L; 23-MAY-01.
DR InterPro: IPR002186; Neocarzinostat.
DR Pfam: PF00960; Neocarzinostat; 1.
DR ProDom: PD012709; Neocarzinostat; 1.
KW Antibiotic; DNA-binding; Signal; 3D-structure.
FT SIGNAL 1 33
FT CHAIN 34 143 ANTITUMOR ANTIBIOTIC C-1027 APOPROTEIN.
FT DISULFID 69 78
FT DISULFID 119 124
SQ SEQUENCE 143 AA; 13873 MW; 8CA3E57FA5DA842A CRC64;

Query Match 75.0%; Score 18; DB 1; Length 143;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
Db 119 CATAAC 124

Search completed: October 1, 2003, 10:42:38
Job time : 11.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:38:06 ; Search time 16.6667 Seconds
(without alignments)
20.309 Million cell updates/sec

Title: US-10-057-890A-4
Perfect score: 24
Sequence: 1 XXXXXC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	18	75.0	9	1	US-08-107-669D-12
2	18	75.0	9	1	US-08-472-788A-12
3	18	75.0	9	2	US-08-477-531B-12
4	18	75.0	9	2	US-08-082-842A-12
5	18	75.0	20	3	US-08-476-509B-47
6	18	75.0	23	3	US-08-084-848A-9
7	18	75.0	23	1	US-08-458-499-9
8	18	75.0	24	2	US-09-105-715-1
9	18	75.0	24	4	US-08-900-230-24
10	18	75.0	24	4	US-08-469-260A-176
11	18	75.0	24	4	US-08-488-446-176
12	18	75.0	24	4	US-08-467-344A-176
13	18	75.0	25	4	US-08-900-230-41
14	18	75.0	26	3	US-09-062-440-12
15	18	75.0	26	4	US-09-712-495-12
16	18	75.0	40	4	US-08-900-230-49
17	18	75.0	43	3	US-08-851-843A-179
18	18	75.0	43	3	US-08-974-549A-298
19	18	75.0	43	3	US-08-854-050-179
20	18	75.0	43	4	US-09-430-323-179
21	18	75.0	43	1	US-08-451-947-98
22	18	75.0	45	3	US-08-424-826A-98
23	18	75.0	45	3	US-08-928-694-98
24	18	75.0	45	4	US-08-900-230-10
25	18	75.0	45	4	US-08-900-230-16
26	18	75.0	45	4	US-08-900-230-39
27	18	75.0	45	4	US-08-900-230-51

28	18	75.0	45	4	US-08-450-842-98	Sequence 98, Appl
29	18	75.0	45	4	US-08-451-390-98	Sequence 98, Appl
30	18	75.0	45	5	PCT-US91-06950-98	Sequence 98, Appl
31	18	75.0	52	2	US-08-799-173A-8	Sequence 8, Appl
32	18	75.0	59	1	US-07-862-021B-15	Sequence 15, Appl
33	18	75.0	59	5	PCT-US93-03164-15	Sequence 15, Appl
34	18	75.0	87	4	US-09-252-991A-25071	Sequence 25071, A
35	18	75.0	93	4	US-09-252-991A-28273	Sequence 28273, A
36	18	75.0	100	4	US-09-252-991A-24179	Sequence 24179, A
37	18	75.0	110	4	US-09-252-991A-18800	Sequence 18800, A
38	18	75.0	111	4	US-09-252-991A-25806	Sequence 25806, A
39	18	75.0	117	1	US-07-942-245-18	Sequence 18, Appl
40	18	75.0	117	4	US-09-252-991A-20556	Sequence 20556, A
41	18	75.0	118	2	US-08-443-639-13	Sequence 13, Appl
42	18	75.0	118	3	US-08-577-483-13	Sequence 13, Appl
43	18	75.0	123	4	US-09-252-991A-30755	Sequence 30755, A
44	18	75.0	125	1	US-07-942-245-33	Sequence 33, Appl
45	18	75.0	126	1	US-08-478-039-95	Sequence 95, Appl

ALIGNMENTS

RESULT 1
US-08-107-669D-12
Sequence 12, Application US/08107669D
Patent No. 5766886
GENERAL INFORMATION:
APPLICANT: Studnicka, Gary M.
TITLE OF INVENTION: Modified Antibody Variable Domains (as amended)
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,669D
FILING DATE: 13-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10906
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,464
FILING DATE: 13-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Michele A. Cimbala
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0610.1000001/MAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/371-2600
TELEFAX: 202/371-2540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-107-669D-12

Query Match 75.0%; Score 18; DB 1; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 4; Indels 0;
Gaps 0;
Qy 3 CXXXXC 8

Db 4 CSSASC 9

RESULT 2
US-08-472-788A-12
; Sequence 12, Application US/08472788A
; Patent No. 5770196
; GENERAL INFORMATION:
; APPLICANT: Studnicka, Gary M.
; TITLE OF INVENTION: Modified Antibody Variable Domains
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,788A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/082,842
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10906
; FILING DATE: 14-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0610.1000003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/371-2540
; TELEFAX: 202/371-2540
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-788A-12

Query Match 75.0%; Score 18; DB 1; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 4 CSSASC 9

RESULT 3
US-08-477-531B-12
; Sequence 12, Application US/08477531B
; Patent No. 582123
; GENERAL INFORMATION:
; APPLICANT: Studnicka, Gary M.
; TITLE OF INVENTION: Modified Antibody Variable Domains (as amended)
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600

CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,531B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/107,669
FILING DATE: 13-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10906
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,464
FILING DATE: 13-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Michele A. Cimbala
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0610.1000004/MAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/371-2600
TELEFAX: 202/371-2540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-531B-12

Query Match 75.0%; Score 18; DB 2; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
DB 4 CSSASC 9

RESULT 4
US-08-082-842A-12
; Sequence 12, Application US/08082842A
; Patent No. 5869619
; GENERAL INFORMATION:
; APPLICANT: Studnicka, Gary M.
; TITLE OF INVENTION: Modified Antibody Variable Domains
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,842A
; FILING DATE: 23-JUN-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10906
; FILING DATE: 14-DEC-1992

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/808.464
; FILING DATE: 13-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0610.1000002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/371-2600
; TELEFAX: 202/371-2540
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-082-842A-12

Query Match 75.0%; Score 18; DB 2; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.5e+05;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
Db 4 CSSASC 9

RESULT 5
US-08-476-509B-47
; Sequence 47, Application US/08476509B
; Patent No. 6034212
; GENERAL INFORMATION:
; APPLICANT: SUDOL, MARIUS
; APPLICANT: PEER, BORK
; APPLICANT: HENRY, CHEN
; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
; TITLE OF INVENTION: SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476.509B
; FILING DATE: 01-DEC-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-101 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "PRIMER"
```

```
; HYPOTHETICAL: NO
US-08-476-509B-47

Query Match 75.0%; Score 18; DB 3; Length 20;
Best Local Similarity 33.3%; Pred. No. 5.2e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
Db 14 CTAATC 19

RESULT 6
US-08-084-848A-9
; Sequence 9, Application US/08084848A
; Patent No. 5432155
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Rivier, Jean E. F.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Abogadie, Fe
; APPLICANT: Hopkins, Chris E.
; APPLICANT: Dykert, John
; APPLICANT: Torres, Josep L.
; TITLE OF INVENTION: Conotoxins I
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 South LaSalle Street, Suite 900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/084.848A
; FILING DATE: June 29, 1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Schumann, James J.
; REGISTRATION NUMBER: 20856
; REFERENCE/DOCKET NUMBER: 52511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)552-1311
; TELEFAX: (619)552-0095
; TELEX: 20 6566 PATLAW CGO
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-084-848A-9

Query Match 75.0%; Score 18; DB 1; Length 23;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
Db 16 CSSASC 21

RESULT 7
US-08-458-499-9
; Sequence 9, Application US/08458499
; Patent No. 5700778
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
```

APPLICANT: Rivier, Jean E. F.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Abogadie, Fe
APPLICANT: Hopkins, Chris E.
APPLICANT: Dykert, John
APPLICANT: Torres, Josep L.
TITLE OF INVENTION: Conotoxins I
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fitch, Even, Tabin & Flannery
STREET: 135 South LaSalle Street, Suite 900
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,499
FILING DATE: 02-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/084,848
FILING DATE: June 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Schumann, James J.
REGISTRATION NUMBER: 20856
REFERENCE/DOCKET NUMBER: 52511
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)552-1311
TELEFAX: (619)552-0095
TELEX: 20 6566 PATLAW CGO
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-458-499-9

Query Match 75.0%; Score 18; DB 1; Length 23;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
Db 16 CSSASC 21

RESULT 8
US-09-105-715-1
Sequence 1, Application US/09105715A
Patent No. 5969096
GENERAL INFORMATION:
APPLICANT: Shoon, Ki-Joon
APPLICANT: Gray, William R.
APPLICANT: Dykert, John
APPLICANT: Yoshikami, Doju
APPLICANT: Watkins, Maren
APPLICANT: Hillyard, David R.
APPLICANT: Rivier, Jean E.F.
APPLICANT: Olivera, Baldomero M.
TITLE OF INVENTION: CONOTOXIN PEPTIDES
FILE REFERENCE: 57823
CURRENT APPLICATION NUMBER: US/09/105,715A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 24

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Conotoxin
OTHER INFORMATION: Peptide Analog
FEATURE:
NAME/KEY: VARIANT
LOCATION: (2)
OTHER INFORMATION: Pro or 4Hyp
FEATURE:
NAME/KEY: VARIANT
LOCATION: (3)
OTHER INFORMATION: Pro or 4Hyp
FEATURE:
NAME/KEY: VARIANT
LOCATION: (14)
OTHER INFORMATION: Pro or 4Hyp
FEATURE:
NAME/KEY: VARIANT
LOCATION: (24)
OTHER INFORMATION: Arg or Gly
US-09-105-715-1

Query Match 75.0%; Score 18; DB 2; Length 24;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
Db 16 CSSASC 21

RESULT 9
US-08-900-230-24
Sequence 24, Application US/08900230
Patent No. 6329197
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A.
TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of The Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 11036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,230
FILING DATE: 23-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-08-900-230-24

Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
| |
Db 16 CATTTC 21

RESULT 10

US-08-469-260A-176
; Sequence 176, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,260A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 176:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-260A-176

Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
| |
Db 8 CSTTAC 13

RESULT 11

US-08-488-446-176
; Sequence 176, Application US/08488446
; Patent No. 6558898
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 176:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-446-176

Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
| |
Db 8 CSTTAC 13

RESULT 12

US-08-467-344A-176
; Sequence 176, Application US/08467344A
; Patent No. 6586568
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER

SHERI L. BUIJK
ISA K. MUSAHWAR
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
REAGENTS AND METHODS FOR THEIR USE
NUMBER OF SEQUENCES: 716
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,344A
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION NUMBER: 08/424,550
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5527.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE DESCRIPTION: SEQ ID NO: 176:
US-08-467-344A-176
Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 33.3%; Pred. NO. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 CXXXXC 8
DB 8 CSTTAC 13
RESULT 13
US-08-900-230-41
; Sequence 41, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; APPLICATION NUMBER: US/08/900,230
; FILING DATE: 23-JUL-1997

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-900-230-41
Query Match 75.0%; Score 18; DB 4; Length 25;
Best Local Similarity 33.3%; Pred. NO. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 CXXXXC 8
DB 18 CATTTC 23
RESULT 14
US-09-062-440-12
; Sequence 12, Application US/09062440
; Patent No. 6207422
; GENERAL INFORMATION:
; APPLICANT: Brown, Arthur M.
; APPLICANT: Yang, Qing
; APPLICANT: Wible, Barbara A.
; TITLE OF INVENTION: A Protein that Enhances Expression of
; TITLE OF INVENTION: Potassium Channels on Cell Surfaces
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter & Griswold LLP
; STREET: 1400 McDonald Investment Center, 800 Superior
; STREET: Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: US
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,440
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Docherty, Pamela A.
; REGISTRATION NUMBER: 40591
; REFERENCE/DOCKET NUMBER: 22884/04004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 2166228200
; TELEFAX: 2162410816
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-062-440-12
Query Match 75.0%; Score 18; DB 3; Length 26;

```

Best Local Similarity 33.3%; Pred. NO. 5.4e-02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXC 8
Db 16 CTTTAC 21

RESULT 15
US-09-712-495-12
; Sequence 12, Application US/09712495
; Patent No. 6391561
; GENERAL INFORMATION:
; APPLICANT: Brown, Arthur
; APPLICANT: Wibler, Barbara
; APPLICANT: Yang, Qing
; TITLE OF INVENTION: Protein That Enhances Expression of Potassium Channels on Cell Sur
; TITLE OF INVENTION: and Nucleic Acids That Encode The Same
; FILE REFERENCE: 22884/04046
; CURRENT APPLICATION NUMBER: US/09/712,495
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 09/062,440
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-712-495-12

Query Match 75.0%; Score 18; DB 4; Length 26;
Best Local Similarity 33.3%; Pred. NO. 5.4e-02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXC 8
Db 16 CTTTAC 21

Search completed: October 1, 2003, 10:49:10
Job time : 17.6667 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:42:07 : Search time 21.3333 Seconds
(without alignments)
59.330 Million cell updates/sec

Title: US-10-057-890a-4
Perfect score: 24
Sequence: 1 XXXXXC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : PublishedApplications_AA.*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	75.0	18	11	US-09-880-748-2770
2	18	75.0	23	11	US-09-833-041-73
3	18	75.0	24	8	US-08-424-550B-176
4	18	75.0	24	15	US-10-072-602B-188
5	18	75.0	24	15	US-10-072-602B-586
6	18	75.0	26	15	US-10-075-074-12
7	18	75.0	30	12	US-10-083-259-143
8	18	75.0	31	15	US-10-072-602B-223
9	18	75.0	31	15	US-10-072-602B-578
10	18	75.0	32	9	US-09-864-761-42539
11	18	75.0	33	9	US-09-864-761-34438
12	18	75.0	33	9	US-09-864-761-35547
13	18	75.0	37	12	US-10-146-984A-2
14	18	75.0	37	12	US-10-289-660-74
15	18	75.0	37	15	US-10-133-128-74

16	18	75.0	37	15	US-10-008-277B-2	Sequence 2, Appli
17	18	75.0	41	12	US-10-063-685-170	Sequence 170, App
18	18	75.0	41	15	US-10-184-644-612	Sequence 612, App
19	18	75.0	41	15	US-10-184-634-612	Sequence 612, App
20	18	75.0	43	10	US-09-843-676-179	Sequence 179, App
21	18	75.0	43	11	US-09-438-486-179	Sequence 179, App
22	18	75.0	43	15	US-10-053-758-179	Sequence 179, App
23	18	75.0	43	15	US-10-054-295-179	Sequence 179, App
24	18	75.0	43	15	US-10-054-611-179	Sequence 179, App
25	18	75.0	45	8	US-08-450-842-98	Sequence 98, Appl
26	18	75.0	56	11	US-09-986-480-300	Sequence 300, App
27	18	75.0	71	12	US-10-269-806-45	Sequence 45, Appl
28	18	75.0	71	12	US-10-269-806-53	Sequence 53, Appl
29	18	75.0	71	12	US-10-269-806-57	Sequence 57, Appl
30	18	75.0	71	12	US-10-269-806-65	Sequence 65, Appl
31	18	75.0	71	12	US-10-269-806-77	Sequence 77, Appl
32	18	75.0	71	12	US-10-269-806-78	Sequence 78, Appl
33	18	75.0	71	12	US-10-269-806-105	Sequence 105, App
34	18	75.0	71	12	US-10-269-806-113	Sequence 113, App
35	18	75.0	71	12	US-10-269-806-117	Sequence 117, App
36	18	75.0	71	15	US-10-011-931-33	Sequence 33, Appl
37	18	75.0	75	9	US-09-758-140-11	Sequence 11, Appl
38	18	75.0	76	15	US-10-072-602B-187	Sequence 187, App
39	18	75.0	77	11	US-09-986-480-444	Sequence 444, App
40	18	75.0	78	15	US-10-072-602B-222	Sequence 222, App
41	18	75.0	78	15	US-10-156-761-7881	Sequence 7881, Ap
42	18	75.0	85	10	US-09-893-737-158	Sequence 158, App
43	18	75.0	88	15	US-10-097-065-469	Sequence 469, App
44	18	75.0	91	11	US-09-764-891-3782	Sequence 3782, App
45	18	75.0	98	9	US-09-864-761-4778	Sequence 4778, A

ALIGNMENTS

RESULT 1
US-09-880-748-2770
; Sequence 2770, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind BlyS
; FILE REFERENCE: PFS23
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2770
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2770

Query Match 75.0%; Score 18; DB 11; Length 18;
Best Local Similarity 33.3%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXC 8

Db 6 CSSTSC 11

RESULT 2

US-09-833-041-73
; Sequence 73, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF345
; CURRENT APPLICATION NUMBER: US/09/833.041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-041-73

Query Match 75.0%; Score 18; DB 11; Length 23;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
|
Db 13 CTTATC 18

RESULT 3
US-08-424-550B-176
; Sequence 176, Application US/08424550B
; Publication No. US20020119447A1
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DARSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUTJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550B
; FILING DATE:
; CLASSIFICATION: 435435
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365

; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 176:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-424-550B-176

Query Match 75.0%; Score 18; DB 8; Length 24;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
|
Db 8 CSTTAC 13

RESULT 4
US-10-072-602B-188
; Sequence 188, Application US/10072602B
; Publication No. US20030109670A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J, Michael
; APPLICANT: Watkins, Maren
; APPLICANT: Garrett, James E.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Grille, Michelle
; APPLICANT: Schoenfeld, Robert M.
; APPLICANT: Walker, Craig
; APPLICANT: Shetty, Reshma
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Cone Snail Peptides
; FILE REFERENCE: 2314-249
; CURRENT APPLICATION NUMBER: US/10/072,602B
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: US 60/267,408
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 188
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Conus purpurascens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(24)
; OTHER INFORMATION: Xaa at residues 2, 3 and 14 is Pro or hydroxy-Pro; Xaa at resi
; OTHER INFORMATION: residues 7 and 13 is Tyr, 125i-Tyr, mono-iodo-Tyr, di-iodo-Tyr
; OTHER INFORMATION: O-sulpho-Tyr or O-phospho-Tyr
US-10-072-602B-188

Query Match 75.0%; Score 18; DB 15; Length 24;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
|
Db 16 CSSASC 21

RESULT 5
US-10-072-602B-586
; Sequence 586, Application US/10072602B
; Publication No. US20030109670A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J, Michael

; APPLICANT: Watkins, Maren
; APPLICANT: Garrett, James E.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Grilley, Michelle
; APPLICANT: Schoenfeld, Robert M.
; APPLICANT: Walker, Craig
; APPLICANT: Shetty, Reshma
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Cone Snail Peptides
; FILE REFERENCE: 2314-249
; CURRENT APPLICATION NUMBER: US/10/072,602B
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: US 60/267,408
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 586
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Conus purpurascens
US-10-072-602B-586

Query Match 75.0%; Score 18; DB 15; Length 24;
Best Local Similarity 33.3%; Pred. No. 5.3e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
| |
Db 16 CSSASC 21

RESULT 6
US-10-075-074-12
; Sequence 12, Application US/10075074
; Publication No. US20030049646A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Arthur
; APPLICANT: Wible, Barbara
; APPLICANT: Yang, Qing
; TITLE OF INVENTION: Protein That Enhances Expression of Potassium Channels on Cell Surfaces
; FILE REFERENCE: 22884/04066
; CURRENT APPLICATION NUMBER: US/10/075,074
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 09/062,440
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/712,495
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-075-074-12

Query Match 75.0%; Score 18; DB 15; Length 26;
Best Local Similarity 33.3%; Pred. No. 5.4e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
| |
Db 16 CTTTAC 21

RESULT 7
US-10-083-259-143
; Sequence 143, Application US/10083259
; Publication No. US20030176325A1
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Peter E.
; APPLICANT: Glierzman, Birgit Kjeldgaard
; APPLICANT: Good, Liam

; APPLICANT: Hansen, Henrik Frydenlund
; APPLICANT: Beck, Frederik
; APPLICANT: Malik, Leila
; APPLICANT: Schou, Carsten
; APPLICANT: Wissenbach, Margit
; TITLE OF INVENTION: Modified Peptide Nucleic Acid (PNA) Molecules
; FILE REFERENCE: PANT-0301
; CURRENT APPLICATION NUMBER: US/10/083,259
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: 09/689,155
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-083-259-143

Query Match 75.0%; Score 18; DB 12; Length 30;
Best Local Similarity 33.3%; Pred. No. 5.6e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
| |
Db 1 CTTTAC 6

RESULT 8
US-10-072-602B-223
; Sequence 223, Application US/10072602B
; Publication No. US20030109670A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Watkins, Maren
; APPLICANT: Garrett, James E.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Grilley, Michelle
; APPLICANT: Schoenfeld, Robert M.
; APPLICANT: Walker, Craig
; APPLICANT: Shetty, Reshma
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Cone Snail Peptides
; FILE REFERENCE: 2314-249
; CURRENT APPLICATION NUMBER: US/10/072,602B
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: US 60/267,408
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 223
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Conus tessulatus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(31)
; OTHER INFORMATION: Xaa at residues 4, 7 and 22 is Glu or gamma-carboxy-Glu; Xaa
; OTHER INFORMATION: Residues 16 and 26 is pro or hydroxy-pro; Xaa at residue 30 is
; OTHER INFORMATION: (D or L) or bromo-Trp (D or L); Xaa at residue 24 is Tyr, 125
; OTHER INFORMATION: Tyr, mono-iodo-Tyr, di-iodo-Tyr, O-sulpho-Tyr or O-phospho-Ty
US-10-072-602B-223

Query Match 75.0%; Score 18; DB 15; Length 31;
Best Local Similarity 33.3%; Pred. No. 5.6e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8

```

Db          9 CSSTAC 14

RESULT 9
US-10-072-602B-578
; Sequence 578, Application US/10072602B
; Publiation No. US20030109670A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Watkins, Maren
; APPLICANT: Garrett, James E.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Grilley, Michelle
; APPLICANT: Schoenfeld, Robert M.
; APPLICANT: Walker, Craig
; APPLICANT: Shetty, Reshma
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: Cone Snail Peptides
; FILE REFERENCE: 2314-249
; CURRENT APPLICATION NUMBER: US/10/072.602B
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: US 60/267,408
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 538
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 578
; TYPE: PRT
; LENGTH: 31
; ORGANISM: Conus tessulatus
US-10-072-602B-578

Query Match          75.0%; Score 18; DB 15; Length 31;
Best Local Similarity 33.3%; Pred. No. 5.6e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy          3 CXXXXC 8
Db          9 CSSTAC 14

RESULT 10
US-09-864-761-42539
; Sequence 42539, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664

US-09-864-761-42539
; Sequence 42539, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42539
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157996.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
; OTHER INFORMATION: EST_HUMAN HIT: BF340389.1, EVALUE 1.50e+00
; OTHER INFORMATION: SWISSPROT HIT: Q43135, EVALUE 3.80e+00
US-09-864-761-42539

Query Match          75.0%; Score 18; DB 9; Length 32;
Best Local Similarity 33.3%; Pred. No. 5.6e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy          3 CXXXXC 8
Db          12 CSSTAC 17

RESULT 11
US-09-864-761-34438
; Sequence 34438, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27

```

; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34438
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007405.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EST_HUMAN HIT: BE614823.1, EVALUE 8.80e-01
; OTHER INFORMATION: SWISSPROT HIT: Q11131, EVALUE 9.50e-01
US-09-864-761-34438

Query Match 75.0%; Score 18; DB 9; Length 33;
Best Local Similarity 33.3%; Pred. No. 5.7e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
| |
Db 16 CTSTSC 21

RESULT 12
US-09-864-761-35547
; Sequence 35547, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35547
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010092.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: EST_HUMAN HIT: BE614823.1, EVALUE 8.80e-01
; OTHER INFORMATION: SWISSPROT HIT: Q11131, EVALUE 9.50e-01
US-09-864-761-35547

Query Match 75.0%; Score 18; DB 9; Length 33;
Best Local Similarity 33.3%; Pred. No. 5.7e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CXXXXC 8
| |
Db 16 CTSTSC 21

RESULT 13
US-10-146-984A-2
; Sequence 2, Application US/10146984A
; Publication No. US20030144309A1
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Moon, Young Choon
; TITLE OF INVENTION: Inhibitors of Src and Other Protein Kinases

; FILE REFERENCE: VPI-01-111
; CURRENT APPLICATION NUMBER: US/10/146.984A
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: US 60/291,340
; PRIOR FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-984A-2

Query Match 75.0%; Score 18; DB 12; Length 37;
Best Local Similarity 33.3%; Pred. No. 5.8e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
Db 23 CATTAC 28

RESULT 14
US-10-289-660-74
; Sequence 74, Application US/10289660
; Publication No. US20030157561A1
; GENERAL INFORMATION:
; APPLICANT: KOLKMAN, JOOST A.
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: GOVINDARAJAN, SRIDHAR
; TITLE OF INVENTION: COMBINATORIAL LIBRARIES OF MONOMER DOMAINS
; FILE REFERENCE: 0319.510US
; CURRENT APPLICATION NUMBER: US/10/289.660
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/133,128
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: 60/374,107
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/286,823
; PRIOR FILING DATE: 2001-04-26
; NUMBER OF SEQ ID NOS: 244
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-289-660-74

Query Match 75.0%; Score 18; DB 12; Length 37;
Best Local Similarity 33.3%; Pred. No. 5.8e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
Db 8 CSTSSC 13

RESULT 15
US-10-133-128-74
; Sequence 74, Application US/10133128
; Publication No. US20030082630A1
; GENERAL INFORMATION:
; APPLICANT: KOLKMAN, JOOST A.
; APPLICANT: STEMMER, WILLEM P.C.
; TITLE OF INVENTION: COMBINATORIAL LIBRARIES OF MONOMER DOMAINS
; FILE REFERENCE: 0319.410US
; CURRENT APPLICATION NUMBER: US/10/133,128
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: 60/374,107

; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/286,823
; NUMBER OF SEQ ID NOS: 244
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-128-74

Query Match 75.0%; Score 18; DB 15; Length 37;
Best Local Similarity 33.3%; Pred. No. 5.8e+02;
Matches 2; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 CXXXXC 8
Db 8 CSTSSC 13

Search completed: October 1, 2003, 10:50:52
Job time : 22.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:04:56 ; Search time 44.7222 Seconds
(without alignments)
24.844 Million cell updates/sec

Title: US-10-057-890A-5

Perfect score: 21

Sequence: 1 HXXXXX 7

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	16	76.2	47	20	AAW92946	HCNV gp116 protein
2	16	76.2	58	20	AAW92945	HCNV gp116 protein
3	16	76.2	106	21	RAB40682	Human ORFX ORF446
4	16	76.2	106	23	ABP33756	Human ORF2729 prot
5	16	76.2	135	20	AAI29154	Amino acid sequenc
6	16	76.2	170	22	AAI67837	Arabidopsis transcr
7	16	76.2	170	22	AAU93024	Porcine reproducti
8	16	76.2	200	22	AAI66797	Novel human diago
9	16	76.2	213	22	ABG26815	

10	16	76.2	237	22	ABG24349	Novel human diago
11	16	76.2	242	21	AAG27670	Arabidopsis thalia
12	16	76.2	278	21	AAG28823	Arabidopsis thalia
13	16	76.2	278	21	AAG45841	Arabidopsis thalia
14	16	76.2	286	21	AAG27669	Arabidopsis thalia
15	16	76.2	290	22	ABB10353	Human cDNA SEQ ID
16	16	76.2	290	23	ABP66940	Human polyptide
17	16	76.2	293	21	AAG27668	Arabidopsis thalia
18	16	76.2	301	23	ABB49391	Listeria monocytog
19	16	76.2	307	22	ABB60804	Drosophila melanog
20	16	76.2	307	22	AAG98892	E. coli growth and
21	16	76.2	313	23	ABP27106	Streptococcus poly
22	16	76.2	325	21	AAG28822	Arabidopsis thalia
23	16	76.2	325	21	AAG45840	Arabidopsis thalia
24	16	76.2	329	21	AAG22780	Arabidopsis thalia
25	16	76.2	329	21	AAG28821	Arabidopsis thalia
26	16	76.2	329	21	AAG50472	Arabidopsis thalia
27	16	76.2	329	22	AAU67934	Propionibacterium
28	16	76.2	345	21	AAG22779	Propionibacterium
29	16	76.2	345	21	AAG50471	Arabidopsis thalia
30	16	76.2	418	22	ABB64546	Drosophila melanog
31	16	76.2	449	22	ABG00347	Novel human diago
32	16	76.2	465	22	ABB57996	Drosophila melanog
33	16	76.2	467	22	AAU63843	Propionibacterium
34	16	76.2	510	21	AAG39697	Arabidopsis thalia
35	16	76.2	514	22	AAG92115	C glutamicum prote
36	16	76.2	545	22	ABG04537	Novel human diago
37	16	76.2	596	23	ABG61825	Prostate cancer-as
38	16	76.2	660	22	ABB69058	Drosophila melanog
39	16	76.2	764	18	AAI26727	Yeast Mpslp protei
40	16	76.2	764	23	ABP54947	Saccharomyces cere
41	16	76.2	779	22	AAI95562	Human protein sequ
42	16	76.2	869	22	ABG19658	Novel human diago
43	16	76.2	906	8	AAI70289	Sequence encoded b
44	16	76.2	906	10	AAI91047	gb envelope protei
45	16	76.2	907	15	AAI63462	Human cytomagalovi

ALIGNMENTS

RESULT 1

AAW92946

ID AAW92946 standard; protein; 47 AA.

XX AC

XX AAW92946;

XX XX

DT 14-MAY-1999 (first entry)

XX XX

DE HCMV gp116 protein fragment #2.

XX XX

KW Detection: anti-HMCV antibody; protective; screening; human plasma;

KW plasma donor; hyperimmunoglobulin; vaccine; infection; glycoprotein;

KW gp116.

XX XX

OS Human cytomegalovirus.

XX XX

PN DE19756214-C1.

XX XX

PD 25-FEB-1999.

XX XX

PF 17-DEC-1997; 97DE-1056214.

XX XX

PR 17-DEC-1997; 97DE-1056214.

XX XX

PA (BIOT-) BIOTEST AG.

XX XX

PI Hinderer W, Lang D, Rothe M, Vornhagen R;

XX XX

WP1: 1999-133568/12.

DR XX

XX New chimeric human cytomegalovirus (HCMV) polypeptides - comprising

PT homologous gp116 fragments from different strains, used to produce

PT vaccines against HCMV infections
 XX
 PS Claim 9; Page 3; 24pp; German.

XX This invention describes a polypeptide which can be used to detect
 CC anti-HCMV antibodies, especially protective antibodies, preferably by
 CC ELISA. The method can also be used to screen human plasma for such
 CC antibodies and to select plasma donors for the production of HCMV
 CC hyperimmunoglobulin or to produce vaccines against HCMV infection. The
 CC novel polypeptide comprises at least 2 polypeptide fragments, each
 CC having an amino acid sequence of a homologous N-terminal region of
 CC different strains of human cytomegalovirus (HCMV) glycoprotein gp116.
 XX

SQ Sequence 47 AA;

Query Match 76.2%; Score 16; DB 20; Length 47;
 Best Local Similarity 28.6%; Pred. No. 3.2e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXH 7
 |
 Db 15 HTTSAAH 21

RESULT 2
 AAW92945
 ID AAW92945 standard; protein; 58 AA.

XX AAW92945;

XX 14-MAY-1999 (first entry)

XX HCMV gp116 protein fragment.

XX Detection: anti-HCMV antibody; protective; screening; human plasma;
 KW plasma donor; hyperimmunoglobulin; vaccine; infection; glycoprotein;
 KW gp116.

XX Human cytomegalovirus.

XX DB19756214-C1.

XX 25-FEB-1999.

XX 17-DEC-1997; 97DE-1056214.

XX 17-DEC-1997; 97DE-1056214.

XX (BIOT-) BIOTEST AG.

XX Hinderer W, Lang D, Rothe M, Vornhagen R;

XX WPI; 1999-133568/12.

XX New chimeric human cytomegalovirus (HCMV) polypeptides - comprising
 PT homologous gp116 fragments from different strains, used to produce
 PT vaccines against HCMV infections
 XX

XX Claim 6; Page 3; 24pp; German.

XX This invention describes a polypeptide which can be used to detect
 CC anti-HCMV antibodies, especially protective antibodies, preferably by
 CC ELISA. The method can also be used to screen human plasma for such
 CC antibodies and to select plasma donors for the production of HCMV
 CC hyperimmunoglobulin or to produce vaccines against HCMV infection. The
 CC novel polypeptide comprises at least 2 polypeptide fragments, each
 CC having an amino acid sequence of a homologous N-terminal region of
 CC different strains of human cytomegalovirus (HCMV) glycoprotein gp116.
 XX

SQ Sequence 58 AA;

Query Match 76.2%; Score 16; DB 20; Length 58;
 Best Local Similarity 28.6%; Pred. No. 3.8e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 HXXXXXH 7
 |
 Db 4 HATSTH 10

RESULT 3

AAB40682
 ID AAB40682 standard; Protein; 106 AA.

XX AAB40682;

XX 08-FEB-2001 (first entry)

XX Human ORFX ORF446 polypeptide sequence SEQ ID NO:892.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX N-PSDB; AAC74891.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 11; Page 907; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,

CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematous, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 106 AA;
 Query Match 76.2%; Score 16; DB 21; Length 106;
 Best Local Similarity 28.6%; Pred. No. 5.8e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 HXXXXXH 7
 DB 94 HSTAT 100
 RESULT 4
 ABP33756
 ID ABP33756 standard; Protein; 106 AA.
 XX
 AC ABP33756;
 DT 09-JUL-2002 (first entry)
 XX
 XX Human ORF2729 protein, SEQ ID NO:5458.
 XX
 KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US17076.
 XX
 PR 24-MAY-2000; 2000US-206690P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Leach MD, Shimkets RA;
 XX
 DR WPI; 2002-106200/14.
 DR N-PSDB; ABN77782.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 XX
 PS Claim 10; Page 1615; 2508pp; English.
 XX
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies

CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antineoplastic activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 XX
 XX Sequence 106 AA;
 Query Match 76.2%; Score 16; DB 23; Length 106;
 Best Local Similarity 28.6%; Pred. No. 5.8e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 HXXXXXH 7
 DB 94 HSTAT 100
 RESULT 5
 AAY29154
 ID AAY29154 standard; Protein; 135 AA.
 XX
 AC AAY29154;
 DT 25-OCT-1999 (first entry)
 XX
 DE Amino acid sequence of a virulence factor encoded by ORF13779.
 XX
 KW Human pathogen; virulence polypeptide; virulence factor;
 KW pathogenic infection; Pseudomonas aeruginosa infection.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN WO9927129-A1.
 XX
 PD 03-JUN-1999.
 XX
 PF 25-NOV-1998; 98WO-US25247.
 XX
 PR 25-NOV-1997; 97US-0066517.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Miklos S;
 PI Rahme IG, Tan M, Tsongalis J;
 XX
 DR WPI; 1999-357851/30.
 XX
 PT Virulence factors useful in developing disease treatments
 XX
 PS Disclosure; Fig 4; 228pp; English.
 XX

CC The present sequence represents a Pseudomonas aeruginosa polypeptide
 CC sequence. P. aeruginosa is an opportunistic human pathogen present in
 CC soil water and plants. The specification describes virulence polypeptides
 CC and nucleic acid sequence encoding such polypeptides. These sequences
 CC can be used to identify a compound which is capable of decreasing the
 CC expression of a pathogenic virulence factor. Compounds that inhibit
 CC the expression or activity of virulence factor polypeptides can be
 CC used to treat pathogenic infections, especially where the infection
 CC is a P. aeruginosa infection.
 CC note: the sequences given in the specification were poorly legible, and
 CC in some instances assumptions were made as to the identity of the
 CC residue; it is therefore possible that the sequence given below is
 CC not entirely correct.

XX Sequence 135 AA;

Query Match 76.2%; Score 16; DB 20; Length 135;
 Best Local Similarity 28.6%; Pred. No. 6.8e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 HXXXXXH 7
 DB 7 HSSAAH 13

RESULT 6
 AAB67837
 ID AAB67837 standard; Protein; 170 AA.

XX AAB67837;

XX 29-JUN-2001 (first entry)

XX Amino acid sequence of a plant transcription factor G736.

XX Transcription factor; flowering time; transgenic plant; vernalisation;
 KW plant development; plant physiology; flowering.

XX Arabidopsis thaliana.

OS WO200126459-A2.

XX 19-APR-2001.

XX 11-OCT-2000; 2000WO-US28141.

XX 12-OCT-1999; 99US-0159464.

PR 08-NOV-1999; 99US-0164132.

PR 17-NOV-1999; 99US-0166228.

PR 17-APR-2000; 2000US-0197899.

PR 22-AUG-2000; 2000US-0227439.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (RATC/) RATCLIFFE O.

PA (HEAR/) HEARD J.

PA (SAMA/) SAMAHA R.

PA (CREE/) CREELMAN R.

PA (KEDD/) KEDDIE J.

PA (JIAN/) JIANG C.

PA (REUB/) REUBER L.

PA (RIEC/) RIECHMANN J L.

XX Ratcliffe O, Heard J, Samaha R, Creelman R, Keddie J, Jiang C;

PI Reuber L, Riechmann JL;

XX WPI; 2001-266398/27.

XX N-PSDB; AAF80407.

XX New transgenic plant comprises a recombinant polynucleotide encoding a
 PT plant transcription factor polypeptide and has a modified flowering
 PT time or vernalization requirement -
 XX Example 7; Page 92; 108pp; English.

XX The present sequence represents a plant transcription factor protein
 CC which modifies the flowering time of a plant. The polynucleotide
 CC sequence is used to produce transgenic plants which have a modified
 CC flowering time or a modified vernalisation requirement. The
 CC polynucleotides and polypeptides are useful for modifying plant
 CC development, physiology or biochemistry such that the modified plants
 CC have a trait advantage over wild type plants. In particular they are
 CC useful for accelerating, delaying or preventing flowering. The
 CC polynucleotides are also useful as nucleic acid probes and primers.
 CC They may be used to identify proteins that can modify the activity of
 CC the transcription factor.

XX Sequence 170 AA;

Query Match 76.2%; Score 16; DB 22; Length 170;
 Best Local Similarity 28.6%; Pred. No. 8e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 HXXXXXH 7
 DB 143 HAAAAAH 149

RESULT 7

AAU93024

ID AAU93024 standard; Protein; 170 AA.

XX AAU93024;

XX 02-JUL-2002 (first entry)

XX Arabidopsis transcription factor #62.

XX Agriculture; metabolic chemical; environmental stress; drought;
 KW microbial disease resistance; herbicide resistance; seed yield;
 KW fruit yield; growth rate; leaf senescence; flower senescence.
 XX plant; transcription factor; transgenic.

OS Arabidopsis thaliana.

XX WO200215675-A1.

XX 28-FEB-2002.

XX 22-AUG-2001; 2001WO-US26189.

XX 22-AUG-2000; 2000US-227439P.

PR 16-NOV-2000; 2000US-0713994.

PR 16-APR-2001; 2001US-0837944.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PA (PILG/) PILGRIM M.

PA (CREE/) CREELMAN R.

PA (DUBE/) DUBELL A J.

PA (HEAR/) HEARD J.

PA (JIAN/) JIANG C.

PA (KEDD/) KEDDIE J.

PA (ADAM/) ADAM L.

PA (RATC/) RATCLIFF O.

PA (REUB/) REUBER J L.

PA (RIEC/) RIECHMANN J L.

PA (YUGG/) YU G.

PA (PINE/) PINEDA O.

XX Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;

PI Adam L, Ratcliffe O, Reuber JL, Riechmann JL, Yu G, Pineda O;

XX WPI; 2002-292022/33.

DR N-PSDB; ABK65210.

XX An isolated or recombinant polynucleotide used to produce a transgenic

PT plant -

xx PS Claim 40; Page 294; 94lpp; English.

xx CC The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ectopic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemicals, pest tolerance, environmental stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf and flower senescence and many other traits listed in the specification). The present sequence is one of the 232 proteins which are A. thaliana transcription factors.

xx SQ Sequence 170 AA;

Query Match 76.2%; Score 16; DB 23; Length 170;
 Best Local Similarity 28.6%; Pred. No. 8e+03; Mismatches 0; Gaps 0;
 Matches 2; Conservative 0; Indels 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
 | |
 Db 143 HAAAAAH 149

RESULT 8
 AAB66797
 ID AAB66797 standard; Protein: 200 AA.
 AC AAB66797;
 XX
 DT 12-APR-2001 (first entry)
 XX
 DE Porcine reproductive and respiratory syndrome virus ORF #5 protein.
 XX Immunodiagnostic; piezoelectric crystal; frequency infection.
 KW Porcine reproductive and respiratory syndrome virus.
 XX OS
 XX W0200102858-A1.
 PN 11-JAN-2001.
 PD
 XX 04-OCT-1999; 99WO-SG00098.
 PF
 XX 05-JUL-1999; 99SG-0003147.
 PR (MOLE-) INST MOLECULAR AGROBIOLOGY.
 PA (MATE-) INST MATERIALS RES & ENG.
 XX LI SFY, Su X, Kwang J, Low S, Liu W;
 PI WPI: 2001-138176/14.
 DR
 XX Detecting the presence of an infectious agent in animals, comprises
 PT immobilizing viral or bacterial specific antigens on a piezoelectric
 PT crystal and measuring the change in resonance frequency of the crystal

PT -
 XX Example 2; Page 41-42; 45pp; English.
 PS
 XX The present invention relates to an immunodiagnostic test for infectious agents in an organism, involving immobilizing an antigen from the agent or an antibody specific for the antigen on a piezoelectric (Pz) crystal, measuring a resonant frequency of the crystal, contacting the crystal with a biological sample, measuring a resonant frequency of the crystal, and comparing the frequencies.

XX SQ Sequence 200 AA;

Query Match 76.2%; Score 16; DB 22; Length 200;
 Best Local Similarity 28.6%; Pred. No. 9e+03; Mismatches 0; Gaps 0;
 Matches 2; Conservative 0; Indels 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
 | |
 Db 32 HSSSSSH 38

RESULT 9
 ABG26815
 ID ABG26815 standard; Protein: 213 AA.
 XX ABG26815;
 AC ABG26815;
 XX 18-FEB-2002 (first entry)
 DT
 XX Novel human diagnostic protein #26806.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS W0200175067-A2.
 PN 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 PI N-PSDB; AAS91002.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
 PT Claim 20; SEQ ID No 57174; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags to identify expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 213 AA;
 Query Match 76.2%; Score 16; DB 22; Length 213;
 Best Local Similarity 28.6%; Pred. No. 9.4e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
 | |
 Db 162 HAASSTH 168

RESULT 10
 ABG24349
 ID ABG24349 standard; Protein: 237 AA.

XX AC ABG24349;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #24340.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang VT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS88336.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX PS Claim 20; SEQ ID NO 54708; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 237 AA;

Query Match 76.2%; Score 16; DB 22; Length 237;
 Best Local Similarity 28.6%; Pred. No. 1e+04;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
 | |
 Db 113 HSSTASH 119

RESULT 11
 AAG27670
 ID AAG27670 standard; Protein: 242 AA.

XX AC AAG27670;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 32598.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 23-APR-1999; 99US-0130891.

XX PR 28-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 06-MAY-1999; 99US-0132487.

XX PR 07-MAY-1999; 99US-0132863.

XX PR 11-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

XX PR 18-MAY-1999; 99US-0134768.

XX PR 19-MAY-1999; 99US-0134941.

XX PR 20-MAY-1999; 99US-0135124.

XX PR 21-MAY-1999; 99US-0135353.

XX PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.


```
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.2%; Score 16; DB 21; Length 242;
Best Local Similarity 28.6%; Pred. No. 1e-04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXH 7
Db 219 HSSSSSH 225

RESULT 12
AAG28823
ID AAG28823 standard; Protein; 278 AA.
XX
AC AAG28823;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 34185.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
```

```
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150984.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155386.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.2%; Score 16; DB 21; Length 278;
Best Local Similarity 28.6%; Pred. No. 1.le+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

QY 1 HXXXXXX 7

DB 243 HAATSSH 249

RESULT 13

AAG45841

ID AAG45841 standard; Protein; 278 AA.

XX AAG45841;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 57601.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

OS EP1033405-A2.

PN 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.2%; Score 16; DB 21; Length 278;
Best Local Similarity 28.6%;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 HXXXXXH 7

Db 243 HAATSSH 249

RESULT 14
AAG27669

ID AAC27669 standard; Protein; 286 AA.
XX AC AAG27669;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 32597.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135144.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139432.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 14-SEP-2000; 2000US-233065P.
PR 21-SEP-2000; 2000US-234221P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 25-SEP-2000; 2000US-234998P.
PR 26-SEP-2000; 2000US-235484P.
PR 27-SEP-2000; 2000US-235834P.
PR 27-SEP-2000; 2000US-235836P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236803P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 13-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 13-OCT-2000; 2000US-239937P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241221P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241786P.
PR 20-OCT-2000; 2000US-241787P.
PR 20-OCT-2000; 2000US-241808P.
PR 20-OCT-2000; 2000US-241809P.
PR 20-OCT-2000; 2000US-241826P.
PR 01-NOV-2000; 2000US-244617P.
PR 08-NOV-2000; 2000US-246474P.
PR 08-NOV-2000; 2000US-246475P.
PR 08-NOV-2000; 2000US-246476P.
PR 08-NOV-2000; 2000US-246477P.
PR 08-NOV-2000; 2000US-246478P.
PR 08-NOV-2000; 2000US-246523P.
PR 08-NOV-2000; 2000US-246524P.
PR 08-NOV-2000; 2000US-246525P.
PR 08-NOV-2000; 2000US-246526P.
PR 08-NOV-2000; 2000US-246527P.
PR 08-NOV-2000; 2000US-246528P.
PR 08-NOV-2000; 2000US-246532P.
PR 08-NOV-2000; 2000US-246609P.
PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246611P.
PR 08-NOV-2000; 2000US-246613P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.
PR 17-NOV-2000; 2000US-249209P.
PR 17-NOV-2000; 2000US-249210P.
PR 17-NOV-2000; 2000US-249211P.
PR 17-NOV-2000; 2000US-249212P.
PR 17-NOV-2000; 2000US-249213P.
PR 17-NOV-2000; 2000US-249214P.
PR 17-NOV-2000; 2000US-249215P.
PR 17-NOV-2000; 2000US-249216P.
PR 17-NOV-2000; 2000US-249217P.
PR 17-NOV-2000; 2000US-249218P.
PR 17-NOV-2000; 2000US-249244P.
PR 17-NOV-2000; 2000US-249245P.
PR 17-NOV-2000; 2000US-249264P.
PR 17-NOV-2000; 2000US-249265P.
PR 17-NOV-2000; 2000US-249297P.
PR 17-NOV-2000; 2000US-249299P.
PR 17-NOV-2000; 2000US-249300P.
PR 01-DEC-2000; 2000US-250160P.
PR 01-DEC-2000; 2000US-250391P.
PR 05-DEC-2000; 2000US-251030P.
PR 05-DEC-2000; 2000US-251988P.
PR 05-DEC-2000; 2000US-256719P.
PR 06-DEC-2000; 2000US-251479P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.

PR 08-DEC-2000; 2000US-251989P.
PR 08-DEC-2000; 2000US-251990P.
PR 11-DEC-2000; 2000US-254097P.
PR 05-JAN-2001; 2001US-259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-476161/51.
DR N-PSDB; ABA06575.
XX
PT Isolated nucleic acid molecule encoding an inflammation-associated
PT polypeptide is used in preventing, treating or ameliorating a medical
condition -
XX
XX Claim 11; SEQ ID NO: 661; 859pp + Sequence Listing; English.
PS
XX The present invention provides human cDNAs, proteins and related genomic
CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a protein of the invention.
XX
SQ Sequence 290 AA;

Query Match 76.2%; Score 16; DB 22; Length 290;
Best Local Similarity 28.6%; Pred. No. 1.2e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXH 7
|
Db 145 HSSSSSH 151

Search completed: October 1, 2003, 10:41:50
Job time : 45.7222 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:34:26 ; Search time 15.3611 Seconds
(without alignments)
43.824 Million cell updates/sec

Title: US-10-057-890A-5

Perfect score: 21

Sequence: 1 HXXXXXH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	76.2	86	2 S40226	glycoprotein B - h
2	16	76.2	87	2 S40231	glycoprotein B - h
3	16	76.2	87	2 S40233	glycoprotein B - h
4	16	76.2	87	2 S40224	glycoprotein B - h
5	16	76.2	87	2 S40232	glycoprotein B - h
6	16	76.2	87	2 S40220	glycoprotein B - h
7	16	76.2	87	2 S40221	glycoprotein B - h
8	16	76.2	87	2 S40222	glycoprotein B - h
9	16	76.2	87	2 S40223	glycoprotein B - h
10	16	76.2	87	2 S40225	glycoprotein B - h
11	16	76.2	87	2 S40230	glycoprotein B - h
12	16	76.2	87	2 S40234	glycoprotein B - h
13	16	76.2	94	2 AD3459	hypothetical prote
14	16	76.2	170	2 H84752	probable DOF zinc
15	16	76.2	248	1 Q0BE4L	probable glycoprot
16	16	76.2	258	2 H82591	hydroxyacylglutath
17	16	76.2	267	2 B69752	beta-lactamase hom
18	16	76.2	294	2 D96542	probable pirin [im
19	16	76.2	300	2 T34071	hypothetical prote
20	16	76.2	301	2 AB1307	dehydrogenases and
21	16	76.2	301	2 AB1679	dehydrogenases and
22	16	76.2	307	2 F90809	probable virulence
23	16	76.2	307	2 S40270	virulence factor m
24	16	76.2	307	2 B85669	probable virulence
25	16	76.2	307	2 A64850	probable virulence
26	16	76.2	307	2 AB0639	probable virulence
27	16	76.2	329	2 A84529	hypothetical prote
28	16	76.2	333	2 T52355	hypothetical prote
29	16	76.2	391	2 T06224	COR39 protein - wh

30 16 76.2 437 2 S55392 msh pteoin - fruit
31 16 76.2 449 2 I40178 probable glutathio
32 16 76.2 536 2 AB0249 probable oxidoredu
33 16 76.2 604 2 T49307 hypothetical prote
34 16 76.2 764 1 S67561 protein kinase RPK
35 16 76.2 775 2 T45136 WD repeat protein
36 16 76.2 801 2 T01289 probable protein k
37 16 76.2 832 2 T33490 hypothetical prote
38 16 76.2 838 2 E89473 protein F5D2.6 [1
39 16 76.2 906 1 VGBEC1 glycoprotein B pre
40 16 76.2 907 1 VGBETE glycoprotein B pre
41 16 76.2 1151 2 T33777 hypothetical prote
42 16 76.2 1616 2 I37183 gene APXL protein
43 16 76.2 1939 2 A80095 Probable sideropho
44 16 76.2 1993 2 T30902 sodium channel SCA
45 16 76.2 3396 2 T22613 hypothetical prote

ALIGNMENTS

RESULT 1

S40226

glycoprotein B - human cytomegalovirus (Chinese isolate patient 1802) (fragment)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 1802
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40226
R:Shiu, S.-Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycopr
A:Reference number: S40220
A:Accession: S40226
A:Molecule type: DNA
A:Residues: 1-86 <SHI>

A:Cross-references: EMBL:X76152; NID:9438020; PIDN:CAA53750.1; PID:9438021

A:Experimental source: Chinese isolate patient 1802

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein

Query Match 76.2%; Score 16; DB 2; Length 86;

Best Local Similarity 28.6%; Pred. No. 4.1e+02;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXH 7

Db 23 HATSSSTH 29

RESULT 2

S40231

glycoprotein B - human cytomegalovirus (Chinese isolate patient 571) (fragment)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 571
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40231

R:Shiu, S.-Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.

submitted to the EMBL Data Library, November 1993

A:Description: Sequence variation of the amino-terminal antigenic domain of glycopr

A:Reference number: S40220

A:Accession: S40231

A:Molecule type: DNA

A:Residues: 1-87 <SHI>

A:Cross-references: EMBL:X76145; NID:9438030; PIDN:CAA53743.1; PID:9438031

A:Experimental source: Chinese isolate patient 571

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein

Query Match 76.2%; Score 16; DB 2; Length 87;

Best Local Similarity 28.6%; Pred. No. 4.1e+02;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXH 7

```

Db      24 HATSSTH 30

RESULT 3
glycoprotein B - human cytomegalovirus (Chinese isolate patient 936) (fragment)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 936
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40233
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprotein
A:Reference number: S40220
A:Accession: S40233
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76155; NID:g438034; PIDN:CAA53753.1; PID:g438035
A:Experimental source: Chinese isolate patient 936
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match      76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HXXXXXXH 7
Db      35 HTTSAAH 41

RESULT 4
glycoprotein B - human cytomegalovirus (Chinese isolate patient 1613 and others) (fragme
S40224
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese patient 1613; Chinese isolate patient 1838; Chinese isolate patient 4
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40224; S40227; S40229
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprotein
A:Reference number: S40220
A:Accession: S40224
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76151; NID:g438016; PIDN:CAA53749.1; PID:g438017
A:Experimental source: Chinese isolate patient 1613
A:Accession: S40227
A:Molecule type: DNA
A:Residues: 1-87 <SHW>
A:Cross-references: EMBL:X76158; NID:g438022; PIDN:CAA53756.1; PID:g438023
A:Experimental source: Chinese isolate patient 1838
A:Accession: S40229
A:Molecule type: DNA
A:Residues: 1-87 <SHF>
A:Cross-references: EMBL:X76153; NID:g438026; PIDN:CAA53751.1; PID:g438027
A:Experimental source: Chinese isolate patient 445
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match      76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HXXXXXXH 7
Db      35 HTTSAAH 41

RESULT 5
glycoprotein B - human cytomegalovirus (Chinese isolate patient 625 and others) (fragmen
S40232
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 625; Chinese isolate patient 1867
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40232; S40228
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprot
A:Reference number: S40220
A:Accession: S40232
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76146; NID:g438032; PIDN:CAA53744.1; PID:g438033
A:Experimental source: Chinese isolate patient 625
A:Accession: S40228
A:Molecule type: DNA
A:Residues: 1-87 <SHW>
A:Cross-references: EMBL:X76159; NID:g438024; PIDN:CAA53757.1; PID:g438025
A:Experimental source: Chinese isolate patient 1867
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match      76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HXXXXXXH 7
Db      24 HATSSTH 30

RESULT 6
glycoprotein B - human cytomegalovirus (Chinese isolate patient 1318) (fragment)
S40220
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 1318
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprot
A:Reference number: S40220
A:Accession: S40220
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76156; NID:g438008; PIDN:CAA53754.1; PID:g438009
A:Experimental source: Chinese isolate patient 1318
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match      76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HXXXXXXH 7
Db      24 HATSSTH 30

RESULT 7
glycoprotein B - human cytomegalovirus (Chinese isolate patient 1512) (fragment)
S40221
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 1512
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40221
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprot
A:Reference number: S40220
A:Accession: S40221
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76148; NID:g438010; PIDN:CAA53746.1; PID:g438011
```


A:Experimental source: Chinese isolate patient 1512
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match 76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
DB 35 HTTSAAH 41

RESULT 8

S40223
glycoprotein B - human cytomegalovirus (Chinese isolate patient 1566) (fragment)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 1566
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40223
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprotein B
A:Reference number: S40220
A:Accession: S40223
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76149; NID:g438012; PIDN:CAA53747.1; PID:g438013
A:Experimental source: Chinese isolate patient 1566
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match 76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
DB 24 HATSSSTH 30

RESULT 9

S40223
glycoprotein B - human cytomegalovirus (Chinese isolate patient 1567) (fragment)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 1567
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40223
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprotein B
A:Reference number: S40220
A:Accession: S40223
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76150; NID:g438014; PIDN:CAA53748.1; PID:g438015
A:Experimental source: Chinese isolate patient 1567
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match 76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
DB 24 HATSSSTH 30

RESULT 10

S40225
glycoprotein B - human cytomegalovirus (Chinese isolate patient 1727) (fragment)
C:Species: human cytomegalovirus, human herpesvirus 5

A:Variety: Chinese isolate patient 1727
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40225
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprotein B
A:Reference number: S40220
A:Accession: S40225
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76157; NID:g438018; PIDN:CAA53755.1; PID:g438019
A:Experimental source: Chinese isolate patient 1727
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match 76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
DB 24 HATSSSTH 30

RESULT 11

S40230
glycoprotein B - human cytomegalovirus (Chinese isolate patient 453) (fragment)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 453
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40230
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprotein B
A:Reference number: S40220
A:Accession: S40230
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76154; NID:g438028; PIDN:CAA53752.1; PID:g438029
A:Experimental source: Chinese isolate patient 453
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match 76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
DB 35 HTTSAAH 41

RESULT 12

S40234
glycoprotein B - human cytomegalovirus (Chinese isolate patient 947) (fragment)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Variety: Chinese isolate patient 947
C:Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S40234
R:Shiu, S.Y.W.; Chan, K.M.; Lo, S.K.F.; Ip, K.W.Y.; Yuen, K.Y.; Heath, R.B.
submitted to the EMBL Data Library, November 1993
A:Description: Sequence variation of the amino-terminal antigenic domain of glycoprotein B
A:Reference number: S40220
A:Accession: S40234
A:Molecule type: DNA
A:Residues: 1-87 <SHI>
A:Cross-references: EMBL:X76147; NID:g438036; PIDN:CAA53745.1; PID:g438037
A:Experimental source: Chinese isolate patient 947
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match 76.2%; Score 16; DB 2; Length 87;
Best Local Similarity 28.6%; Pred. No. 4.1e+02;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
Db 24 HATSSSTH 30

RESULT 13

AD3459
hypotheical protein BMEI1658 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AD3459
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Muler, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzev, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3459
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-94 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52839.1; PID:gl7983680; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1658
A:Map position: 1

Query Match 76.2%; Score 16; DB 2; Length 94;
Best Local Similarity 28.6%; Pred. No. 4.3e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
Db 45 HAATSTH 51

RESULT 14

H84752
probable DOF zinc finger protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84752
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84752
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-170 <STO>
A:Cross-references: GB:AE002093; NID:g2342734; PIDN:AAB67632.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g34140
A:Map position: 2

Query Match 76.2%; Score 16; DB 2; Length 170;
Best Local Similarity 28.6%; Pred. No. 6.9e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
Db 143 HAAAAAH 149

RESULT 15

Q0BE4L
probable glycoprotein - human herpesvirus 4 (strain B95-8)
C:Species: human herpesvirus 4, Epstein-Barr virus
C>Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 16-Jul-1999
C:Accession: A03780; S33052

R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr vi
A:Reference number: A93065; MUID:85035713; PMID:6092825
A:Accession: A03780
A:Molecule type: DNA
A:Residues: 1-248 <BAN>
A:Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24803.1; PID:gl334911
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.
Nature 310, 207-211, 1984
A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A:Reference number: A03794; MUID:84270567; PMID:6087149
A:Contents: annotation: protein coding region
C:Superfamily: Epstein-Barr virus glycoprotein
C:Keywords: glycoprotein
F:27,45,73,83,92,95,104,116,131,144/Binding site: carbohydrate (Asn) (covalent) #stat

Query Match 76.2%; Score 16; DB 1; Length 248;
Best Local Similarity 28.6%; Pred. No. 9.2e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
Db 169 HTTSTSH 175

Search completed: October 1, 2003, 10:47:50
Job time : 17.3611 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:06:46 ; Search time 8.16667 Seconds
(without alignments)
40.309 Million cell updates/sec

Title: US-10-057-890A-5
Perfect score: 21
Sequence: 1 HXXXXXX 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SWISSProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	76.2	248	1 VGL2_EBV	P03218 Epstein-Barr
2	16	76.2	267	1 VGL2_EBV	P54427 bacillus su
3	16	76.2	307	1 VGL2_EBV	P75931 escherichia
4	16	76.2	307	1 VGL2_EBV	P37168 salmonella
5	16	76.2	310	1 PIR1_ARATH	G91P89 arabidopsis
6	16	76.2	440	1 BRA2_BRAFL	P80492 branchiosto
7	16	76.2	448	1 BRA1_BRAFL	Q17134 branchiosto
8	16	76.2	449	1 GSHR_BURCE	P48639 burkholderi
9	16	76.2	515	1 HNSH_DROME	Q03372 drosophila
10	16	76.2	623	1 NRC2_NEUCR	Q42626 neurospora
11	16	76.2	764	1 MFS1_YEAST	P54199 saccharomyc
12	16	76.2	775	1 POP1_SCHPO	P87060 schizosacch
13	16	76.2	905	1 VGLB_HCNVA	P06473 human cytom
14	16	76.2	907	1 VGLB_HCNVT	P13201 human cytom
15	16	76.2	929	1 ART1_HUMAN	Q92081 h adipocyte
16	16	76.2	930	1 ART1_MOUSE	Q9eqh2 mus musculu
17	16	76.2	930	1 ART1_MOUSE	Q9jjj2 rattus norv
18	16	76.2	1147	1 CARB_HUMAN	G9bx17 homo sapien
19	16	76.2	1616	1 APXL_HUMAN	Q13796 homo sapien
20	16	76.2	1850	1 BA2A_MOUSE	Q91y65 mus musculu
21	16	76.2	1878	1 BA2A_HUMAN	Q9uif9 homo sapien
22	16	76.2	4705	1 FAT2_DROME	Q9vw71 drosophila
23	15	71.4	54	1 K116_HUMAN	Q95650 homo sapien
24	15	71.4	95	1 CFF7_MYCTU	O53693 mycobacteri
25	15	71.4	130	1 IGFL_CAVPO	P17647 cavia porce
26	15	71.4	153	1 IGFL_HUMAN	P16545 sus scrofa
27	15	71.4	153	1 IGFL_HUMAN	P01343 homo sapien
28	15	71.4	167	1 MBP_CAVPO	P25188 cavia porce
29	15	71.4	168	1 MBP_RABIT	P25274 oncotolagus
30	15	71.4	169	1 MBP_BOVIN	P02687 bos taurus
31	15	71.4	171	1 MBP_PIG	P81558 sus scrofa
32	15	71.4	173	1 MBP_CHICK	P15720 gallus gall
33	15	71.4	195	1 IGFB_HUMAN	P05019 homo sapien

34	15	71.4	198	1 SECG_HELPF	Q9z1x2 helicobacte
35	15	71.4	204	1 CYOC_ECOLI	P18402 escherichia
36	15	71.4	218	1 IPVR_ARATH	P21216 arabidopsis
37	15	71.4	220	1 UPAS_RAT	P51573 rattus norv
38	15	71.4	229	1 YH07_STAAM	Q99t16 staphylococ
39	15	71.4	240	1 YDFB_SCHPO	Q10483 schizosacch
40	15	71.4	245	1 YCDX_ECOLI	P75914 escherichia
41	15	71.4	249	1 YS87_CAEEL	Q09383 caenorhabdi
42	15	71.4	250	1 SURE_VIBCH	Q9ku19 vibrio chol
43	15	71.4	250	1 TXRH_AERPE	Q9y910 aeropyrum p
44	15	71.4	251	1 ATPD_PEA	Q02758 pisum sativ
45	15	71.4	253	1 IF46_TOBAC	Q40469 nicotiana t

ALIGNMENTS

RESULT 1
VGL2_EBV
ID VGL2_EBV STANDARD; PRT: 248 AA.
AC P03218;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable membrane glycoprotein.
GN BILF2.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85035713; PubMed=6092825;
RA Bankier A.T., Deininger P.L., Farrell P.J., Barrell B.G.;
RT *Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8
RT Epstein-Barr virus.*;
RL Mol. Biol. Med. 1:21-45(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RT Tufnell P.S., Barrell B.G.;
RT *DNA sequence and expression of the B95-8 Epstein-Barr virus genome.*;
RN Nature 310:207-211(1984).
RP SEQUENCE FROM N.A.
RX MEDLINE=90244363; PubMed=2159529;
RA Mackett M., Conway M.J., Arrand J.R., Haddad R.S.,
RA Hutt-Fletcher L.M.;
RT *Characterization and expression of a glycoprotein encoded by the
RT Epstein-Barr virus BamHI I fragment.*;
RL J. Virol. 64:2545-2552(1990).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01555; CAA24803.1; -
DR EMBL; M37129; AAA45876.1; -
DR PIR; A03780; Q0BE4L.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003096; IG_MHC.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
KW Immunoglobulin domain; Glycoprotein; Late protein; Membrane.

```

FT DOMAIN 19 125 IG-LIKE.
SQ SEQUENCE 248 AA; 27076 MW; C3F33A253B959ADA CMC64;

Query Match 76.2%; Score 16; DB 1; Length 248;
Best Local Similarity 28.6%; Pred. No. 4.9e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 HXXXXXX 7
Db 169 HTTSTSH 175

RESULT 2
YBXI_BACSU STANDARD; PRT; 267 AA.
ID YBXI_BACSU STANDARD; PRT; 267 AA.
AC P54427; O31439;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable beta-lactamase ybxI precursor (EC 3.5.2.6).
GN YBXI OR YBOS.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RA Haga K., Liu H., Yasumoto K., Takahashi H., Yoshikawa H.;
RT "Sequence analysis of the 70kb region between 17 and 23 degree of the
RT Bacillus subtilis chromosome."
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Arevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brullet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Hache J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kjaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilwa A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle R., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassaretti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
RN [3]
RN SEQUENCE OF 129-267 FROM N.A.
RC STRAIN=168 / PY79;
RX MEDLINE=97169156; PubMed=9016963;
RA Schepetov M., Chyu G., Bagyan I., Cutting S.M.;
RT "Characterization of csga, a new member of the forespore-expressed
RT sigmaG-regulon from Bacillus subtilis."

```

Gene 184:133-140(1997).

CC -I- CATALYTIC ACTIVITY: A beta-lactam + H(2)O -> a substituted beta-amino acid.

CC -I- SIMILARITY: BELONGS TO THE CLASS-D BETA-LACTAMASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AB006424; BAA33106.1; -

DR EMBL: Z99105; CAB12003.1; -

DR EMBL: X92859; CAA63445.1; -

DR PIR: B69752; B69752.

DR HSP: P14489; 1E3U.

DR Subtilist; BG11507; ybxI.

DR InterPro: IPR002137; Beta_lactamase_D.

DR PROSITE: PS00337; BETA-LACTAMASE_D; FALSE_NEG.

KW Hypothetical protein; Hydrolase; Antibiotic resistance; Signal;

KW Complete proteome.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 267 PROBABLE BETA-LACTAMASE YBXI.

FT ACT_SITE 76 76 BY SIMILARITY.

FT BINDING 214 216 SUBSTRATE (BY SIMILARITY).

FT CONFLICT 129 137 YQALRDIG -> IISISERY (IN REF. 3).

FT CONFLICT 169 169 T -> A (IN REF. 3).

SQ SEQUENCE 267 AA; 30644 MW; BBE3E698CD46E6ED CRC64;

Query Match 76.2%; Score 16; DB 1; Length 267;

Best Local Similarity 28.6%; Pred. No. 5.2e+02;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 HXXXXXX 7

Db 22 HAASAH 28

RESULT 3

YBXI_BACSU STANDARD; PRT; 307 AA.

AC P75931; (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Virulence factor mvim homolog.

CN MVIM OR B1068.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=97061202; PubMed=8905232;

RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Itoh T., Kajihara M., Kanai K., Kashimoto K.,

RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,

RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,

RA Sampaio G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,

RA Yano M., Horiuchi T.;

```

RT      *A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT      corresponding to the 12.7-28.0 min region on the linkage map.*;
RL      DNA Res. 3:137-155(1996).
CC      -1- SIMILARITY: TO S.TYPHIMURIUM MVIM.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; AE000207; AAC74152.1; -.
DR      EMBL; D90743; BA35876.1; -.
DR      PIR; A64850; A64850.
DR      EcoGene; EG13879; mvim.
DR      InterPro; IPR000683; GFO_IDH_MoCA.
DR      InterPro; IPR004104; GFO_IDH_MoCA_C.
DR      Pfam; PF01408; GFO_IDH_MoCA; 1.
DR      Pfam; PF02894; GFO_IDH_MoCA_C; 1.
DR      Virulence; Complete proteome.
KW      Virulence: Complete proteome.
SQ      SEQUENCE 307 AA; 33681 MW; F8BB622AE0C5C92C CRC64;

Query Match      76.2%   Score 16; DB 1; Length 307;
Best Local Similarity 28.6%   Pred. No. 5.8e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1 HXXXXXX 7
Db      70 HSTASH 76

RESULT 4
MVIM_SALTY
ID      MVIM_SALTY STANDARD; PRT; 307 AA.
AC      P37168;
DT      01-OCT-1994 (Rel. 30, Created)
DT      01-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Virulence factor mvim.
DE      MVIM OR STM1169.
OS      Salmonella typhimurium.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxID=602;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LT2 / KK1004;
RA      van Slooten J.-C., Okada T., Kutsukake K., Pechere J.-C., Harayama S.;
RL      Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LT2;
RA      Kutsukake K., Okada T., Yokoseki T., Iino T.;
RX      MEDLINE=94259301; PubMed=8200538;
RT      *Sequence analysis of the flgA gene and its adjacent region in
RT      Salmonella typhimurium, and identification of another flagellar gene,
RT      flgN.*;
RL      Gene 143:49-54(1994).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX      MEDLINE=21534948; PubMed=11677609;
RA      McClelland M., Sanderson K.E., Spleth J., Clifton S.W., Latreille P.,
RA      Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA      Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA      Ryan E., Sun H., Florea L., Miller W., Stoneking I., Nhan M.,
RA      Waterston R., Wilson R.K.;
RT      *Complete genome sequence of Salmonella enterica serovar Typhimurium
RT      LT2.*;
RL      Nature 413:852-856(2001).
CC      -1- SIMILARITY: TO E.COLI MVIM.
-----

```

```

CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; 226133; CA811133.1; -.
DR      EMBL; D25292; BAA04981.1; -.
DR      EMBL; AE008750; AAL20099.1; -.
DR      PIR; S40270; S40270.
DR      StyGene; SG10239; mvim.
DR      InterPro; IPR000683; GFO_IDH_MoCA.
DR      InterPro; IPR004104; GFO_IDH_MoCA_C.
DR      Pfam; PF01408; GFO_IDH_MoCA; 1.
DR      Pfam; PF02894; GFO_IDH_MoCA_C; 1.
KW      Virulence; Complete proteome.
FT      CONFLICT 163 163 D -> H (IN REF. 1).
FT      CONFLICT 194 194 E -> G (IN REF. 1).
SQ      SEQUENCE 307 AA; 34008 MW; 7CD105D8DFC8B14F CRC64;

Query Match      76.2%   Score 16; DB 1; Length 307;
Best Local Similarity 28.6%   Pred. No. 5.8e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      1 HXXXXXX 7
Db      70 HSTASH 76

RESULT 5
PIRL_ARATH
ID      PIRL_ARATH STANDARD; PRT; 310 AA.
AC      Q9LPS9; Q9C6P9;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Pirin-like protein Atlg50590.
GN      ATLG50590 OR Flil2.9 OR F17J6.11.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsids.
CX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Columbia;
RX      MEDLINE=21016719; PubMed=11130712;
RA      Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA      White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA      Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA      Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA      Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA      Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA      Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA      Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA      Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA      Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA      Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA      Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA      Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA      Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA      Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA      Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT      *Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT      thaliana.*;
RL      Nature 408:816-820(2000).
CC      -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE PIRIN FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AC012561; AAF87876.1; -;
 DR EMBL; AC079279; AAG51187.1; -;
 DR PIR; D96542; D96542.
 DR InterPro; IPR003829; DUF209.
 DR Pfam; PF02678; Pirin; 1.
 DR Nuclear protein.
 KW CONFLICT 68 83 MISSING (IN REF. 1; AAG51187).
 SQ SEQUENCE 310 AA; 34547 MW; 6C07034305B31782 CRC64;

Query Match 76.2%; Score 16; DB 1; Length 310;
 Best Local Similarity 28.6%; Pred. No. 5.9e+02;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXX 7
 Db 221 HSAHAH 227

RESULT 6
 BRA2_BRAFL STANDARD; PRT; 440 AA.
 ID BRA2_BRAFL
 AC P80492;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Brachyury protein homolog 2 (AmBra-2).
 OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
 GN Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OX NCBI_TaxID=7739;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96125169; PubMed=8575328;
 RA Holland P.W.H., Koschorz B., Holland L.Z., Herrmann B.G.;
 RT "Conservation of Brachyury (T) genes in amphioxus and vertebrates:
 development and evolutionary implications.";
 RL Development 121:4283-4291(1995).
 CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
 CC REQUIRED FOR MESODERM FORMATION AND DIFFERENTIATION
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED IN THE 10 H GASTRULA AROUND
 CC THE BLASTOPORE. IN LATER STAGES OF GASTRULATION AND EARLY
 CC NEURALIZATION, EXPRESSED IN POSTERIOR MESODERM AND THE NOTOCHORD
 CC WHERE EXPRESSION CONTINUES AT LEAST UNTIL SWIMMING LARVAL STAGE.
 CC -1- SIMILARITY: Contains 1 T-box domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X91903; CAA62999.1; -;
 DR HSP; P24781; IXBR.
 DR InterPro; IPR001699; TF_T-box.
 DR Pfam; PF00907; T-box; 1.
 DR PRINTS; PR00937; TBOX.
 DR SMART; SM00425; TBOX; 1.
 DR PROSITE; PS01283; TBOX_1; 1.
 DR PROSITE; PS01264; TBOX_2; 1.
 DR PROSITE; PS0252; TBOX_3; 1.
 KW Developmental protein; Transcription regulation; DNA-binding;
 KW Nuclear protein; Multigene family.
 FT DNA_BIND 44 215
 T-BOX.
 SQ SEQUENCE 440 AA; 48469 MW; F12EBC417243D3D6 CRC64;

Query Match 76.2%; Score 16; DB 1; Length 440;
 Best Local Similarity 28.6%; Pred. No. 7.7e+02;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXX 7
 Db 332 HTTSTH 338

RESULT 7
 BRA1_BRAFL STANDARD; PRT; 448 AA.
 ID BRA1_BRAFL
 AC Q17134;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Brachyury protein homolog 1 (AmBra-1).
 GN BRA-1.
 OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
 GN Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OX NCBI_TaxID=7739;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Larva;
 MEDLINE=96125169; PubMed=8575328;
 RA Holland P.W.H., Koschorz B., Holland L.Z., Herrmann B.G.;
 RT "Conservation of Brachyury (T) genes in amphioxus and vertebrates:
 development and evolutionary implications.";
 RL Development 121:4283-4291(1995).
 CC -1- FUNCTION: INVOLVED IN THE TRANSCRIPTIONAL REGULATION OF GENES
 CC REQUIRED FOR MESODERM FORMATION AND DIFFERENTIATION
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED IN THE 10 H GASTRULA AROUND
 CC THE BLASTOPORE. IN LATER STAGES OF GASTRULATION AND EARLY
 CC NEURALIZATION, EXPRESSED IN POSTERIOR MESODERM AND THE NOTOCHORD
 CC WHERE EXPRESSION CONTINUES AT LEAST UNTIL SWIMMING LARVAL STAGE.
 CC -1- SIMILARITY: Contains 1 T-box domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X91903; CAA62999.1; -;
 DR HSP; P24781; IXBR.
 DR InterPro; IPR001699; TF_T-box.
 DR Pfam; PF00907; T-box; 1.
 DR PRINTS; PR00937; TBOX.
 DR SMART; SM00425; TBOX; 1.
 DR PROSITE; PS01283; TBOX_1; 1.
 DR PROSITE; PS01264; TBOX_2; 1.
 DR PROSITE; PS0252; TBOX_3; 1.
 KW Developmental protein; Transcription regulation; DNA-binding;
 KW Nuclear protein; Multigene family.
 FT DNA_BIND 54 224
 T-BOX.
 SQ SEQUENCE 448 AA; 49253 MW; 43326D32FD4BC156 CRC64;

Query Match 76.2%; Score 16; DB 1; Length 448;
 Best Local Similarity 28.6%; Pred. No. 7.9e+02;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXX 7
 Db 340 HTTSTH 346

RESULT 8
 GSHR_BURCE STANDARD; PRT; 449 AA.
 ID GSHR_BURCE
 AC P48639;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glutathione reductase (EC 1.8.1.7) (GR) (Grase).
 GN GOR.
 OS Burkholderia cepacia (Pseudomonas cepacia).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

```
CC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AC1100;
RX MEDLINE=95266809; PubMed=7538273;
RA Daubaras D.L., Hersberger C.D., Kitano K., Chakrabarty A.M.;
RT "Sequence analysis of a gene cluster involved in metabolism of 2,4,5-
trichlorophenoxyacetic acid by Burkholderia cepacia AC1100.";
RL Appl. Environ. Microbiol. 61:1279-1289(1995).
CC -!- CATALYTIC ACTIVITY: 2 glutathione + NADP(+) = glutathione
disulfide + NADPH.
CC -!- COFACTOR: FAD.
CC -!- PATHWAY: 2,4,5-trichlorophenoxyacetic acid degradation.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -!- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
OXIDOREDUCTASES CLASS-I.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: U19883; AAC43334.1; -.
DR PIR: I40178; I40178.
DR HSSP: P28593; IAOG.
DR InterPro: IPR001327; FAD_Pyr_redox.
DR InterPro: IPR006324; Glut_reductase.
DR InterPro: IPR000815; Hg_reductase.
DR InterPro: IPR001100; Pyr_redox.
DR InterPro: IPR004099; Pyr_redox_dim.
DR Pfam: PF00070; Pyr_redox; 1.
DR Pfam: PF02852; Pyr_redox_dim; 1.
DR PRINTS: PR00368; FADPNR.
DR PRINTS: PR00945; HGRDTASEI.
DR PRINTS: PR00411; PNDRDTASEI.
DR ProDom: PD000139; FAD_Pyr_redox; 1.
DR TIGRFAMs: TIGR01424; glut_redox; 2; 1.
DR PROSITE: PS00076; PYRIDINE_REDOX_1; 1.
DR Redox-active center; Oxidoreductase; Flavoprotein; FAD; NADP.
FT NP_BIND 7 37 FAD (ADP PART) (PROBABLE).
FT DISULFID 43 48 REDOX-ACTIVE (BY SIMILARITY).
FT NP_BIND 292 302 FAD (FLAVIN PART) (BY SIMILARITY).
FT ACT_SITE 435 435 BY SIMILARITY.
SQ SEQUENCE 449 AA; 47541 MW; 402FCC6E7A8D6720 CRC64;

Query Match 76.2%; Score 16; DB 1; Length 449;
Best Local Similarity 28.6%; Pred. No. 7.9e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
Db 73 HSAATSH 79

RESULT 9
HMSH_DROME STANDARD; PRT; 515 AA.
AC Q03372;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Muscle segmentation homeobox (protein Drop).
GN DR OR MSH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
```

```
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97042058; PubMed=8887329;
RA D'Alessio M., Frasch M.;
RT "msh may play a conserved role in dorsoventral patterning of the
neuroectoderm and mesoderm.";
RL Mech. Dev. 58:217-231(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=98146200; PubMed=9486795;
RA Nose A., Tshiki T., Takeichi M.;
RT "Regional specification of muscle progenitors in Drosophila: the role
of the msh homeobox gene.";
RL Development 125:215-223(1998).
RN [3]
RP SEQUENCE OF 420-480 FROM N.A.
RX MEDLINE=91200674; PubMed=1673109;
RA Holland P.W.H.;
RT "Cloning and evolutionary analysis of msh-like homeobox genes from
mouse, zebrafish and ascidian.";
RL Gene 98:253-257(1991).
CC -!- FUNCTION: PROBABLE MORPHOGENETIC ROLE IN PRONEURAL AND PROMUSCULAR
CLUSTER FORMATION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE MSH HOMEBOX FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: U33319; AAC47329.1; -.
DR EMBL: AF009038; AAB62975.1; -.
DR EMBL: M38582; AAA28611.1; -.
DR PIR: PS0404; PS0404.
DR HSSP: P14653; IB72.
DR TRANSFAC: T03551; -.
DR FlyBase: FBgn0000492; Dr.
DR GO: GO:0007450; P:dorsal/ventral pattern formation, imaginal . . .; IMP.
DR GO: GO:0007419; P:ventral cord development; NAS.
DR GO: GO:0007476; P:wing morphogenesis; IMP.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DOMAIN 38 45 POLY-SER.
FT DOMAIN 94 108 POLY-GLN.
FT DOMAIN 175 179 POLY-PRO.
FT DOMAIN 205 214 POLY-ALA.
FT DOMAIN 225 229 POLY-GLN.
FT DOMAIN 364 371 POLY-ALA.
FT DNA_BIND 421 480 HOMEBOX.
SQ SEQUENCE 515 AA; 54264 MW; AE475EBA29320A0D CRC64;

Query Match 76.2%; Score 16; DB 1; Length 515;
Best Local Similarity 28.6%; Pred. No. 8.8e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
Db 207 HAAAAAH 213

RESULT 10
```

NRC2_NEUCR
ID NRC2_NEUCR STANDARD; PRT; 623 AA.
AC O42626;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine protein kinase nrc-2 (EC 2.7.1.37) (Nonrepressible
DE condensation protein 2).
GN NRC-2 OR BIK1.130.
GN Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OS Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OC NCBL_TaxID=5141;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=98250677; PubMed=9584090;
RX Kothé G.O., Free S.J.;
RA "The isolation and characterization of nrc-1 and nrc-2, two genes
RA encoding protein kinases that control growth and development in
RA Neurospora crassa";
RT Genetics 149:117-130(1998).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=74-OR23-1A / FGSC 987;
RX PubMed=12655011;
RA Mannhaupt G., Montrone C., Haase D., Mewes H.-W., Align V.,
RA Hoheisel J.D., Fartmann B., Nykatura G., Kempken F., Maier J.,
RA Schulte U.;
RT "What's in the genome of a filamentous fungus? Analysis of the
RT Neurospora genome sequence";
RL Nucleic Acids Res. 31:1944-1954(2003).
CC -!- FUNCTION: Controls entry of the cell into the asexual
CC developmental program. Required to repress entry into the
CC condensation program.
CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphorylated protein.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC KIN82 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF034260; AAC21677.1; -;
DR EMBL; AL669998; CAD21180.1; -;
DR HSSP; P05132; IAPM.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Condensation.
FT DOMAIN 242 532 PROTEIN KINASE.
FT NP_BIND 248 256 ATP (BY SIMILARITY).
FT BINDING 271 271 ATP (BY SIMILARITY).
FT ACT_SITE 367 367 BY SIMILARITY.
SQ SEQUENCE 623 AA; 68628 MW; 17B3D2FB73E1ACE7 CRC64;
Query Match 76.2%; Score 16; DB 1; Length 623;
Best Local Similarity 28.6%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 HXXXXX 7
|
|

Db 36 HTTSSSH 42
RESULT 11
MPS1_YEAST STANDARD; PRT; 764 AA.
AC P54139;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine protein kinase MPS1 (EC 2.7.1.37) (Regulatory cell
DE proliferation kinase 1).
GN MPS1 OR RPK1 OR YDL028C OR D2785.
OS Saccharomycetes cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetidae; Saccharomycetaceae; Saccharomycetes.
OX NCBL_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=FL100;
RX MEDLINE=94301294; PubMed=8028580;
RA Poch O., Schwob E., de Fraipont F., Camasses A., Bordonne R.,
RA Martin R.P.;
RT "RPK1, an essential yeast protein kinase involved in the regulation
RT of the onset of mitosis, shows homology to mammalian dual-specificity
RT kinases";
RL Mol. Gen. Genet. 243:641-653(1994).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C;
RX MEDLINE=97197972; PubMed=9046088;
RA Saren A.M., Laamanen P., Lejarcegui J.B., Paulin L.;
RT "The sequence of a 36.7 kb segment on the left arm of chromosome IV
RT from Saccharomycetes cerevisiae reveals 20 non-overlapping open reading
RT frames (ORFs) including SPT4, FAD1, NAM1, RN11, SIR2, NAT1, PRP9,
RT AC12 and MPS1 and 11 new ORFs";
RL Yeast 13:65-71(1997).
RN [3]
RN CHARACTERIZATION, AND MUTAGENESIS OF ASP-580.
RX MEDLINE=95255223; PubMed=7737118;
RA Lauze E., Stoelcker B., Luca F.C., Weiss E., Schutz A.R., Winey M.;
RT "Yeast spindle pole body duplication gene MPS1 encodes an essential
RT dual specificity protein kinase";
RL EMBL J. 14:1855-1863(1995).
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF THE ONSET OF MITOSIS.
CC INVOLVED IN A PATHWAY THAT COORDINATES CELL PROLIFERATION AND
CC DIFFERENTIATION. IMPLICATED IN SPINDLE POLE BODY (SPB)
CC DUPLICATION. DUAL SPECIFICITY KINASE THAT CAN PHOSPHORYLATE
CC SERINE, THREONINE AND TYROSINE RESIDUES.
CC -!- PTM: AUTOPHOSPHORYLATED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L08909; AAA88731.1; -;
DR EMBL; 271781; CAA96461.1; -;
DR EMBL; 274076; CAA98587.1; -;
DR PIR; S67561; S67561.
DR SGD; S0002186; MPS1.
DR GO; GO:0005699; C:kinetochore; IDA.
DR GO; GO:0005816; C:spindle pole body; IDA.
DR GO; GO:0004712; F:protein threonine/tyrosine kinase activity; IDA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.

DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR	PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW	Transferase; Serine/threonine-protein kinase; ATP-binding;
KW	phosphorylation..
FT	DOMAIN 273 282 POLY-SER.
FT	DOMAIN 309 315 POLY-SER.
FT	DOMAIN 440 720 PROTEIN KINASE.
FT	NP_BIND 446 454 ATP (BY SIMILARITY).
FT	N_BINDING 468 468 ATP (BY SIMILARITY).
FT	ACT_SITE 563 563 BY SIMILARITY.
FT	MUTAGEN 580 580 D->A; LOSS OF ACTIVITY.
FT	CONFLICT 146 146 A -> S (IN REF. 1).
FT	CONFLICT 211 213 RRE -> TKR (IN REF. 1).
SO	SEQUENCE 764 AA: 86827 MW: 26817 DEFF8BEB4D CRC64:

Query Match 76.2%; Score 16; DB 1; Length 764;
Best Local Similarity 28.6%; Pred. No. 1.2e+03;
Matches 2; Conservative 0; Mismatches 5; Indels

Qy	1 HXXXXXH 7
Dp	88 HTSTSSH 9.

RESULT 12

POP1_SCHPO	STANDARD:	PRT:	775 AA.
AC	POP1_SCHPO		
AD	P87060; Q9P7P3;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	WD-repeat protein popl.		
DE	POP1 OR SPEC118.01 OR SPBC2G2.18.		
GN	Schizosaccharomyces pombe (fission yeast).		
OS	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;		
OC	Schizosaccharomycetes.		
OX	NCBI_TaxID=4896;		
[1]			
RN	SEQUENCE FROM N.A.		
RP	STRAIN-972;		
RC	MEDLINE-97347242; PubMed-9203581;		
RX	Kominami K., Toda T.;		
RA	*Fission yeast WD-repeat protein popl regulates genome ploidy through		
RT	ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rum1		
RT	and the S-phase Initiator Cdc18.*;		
RL	Genes Dev. 11:1548-1560(1997).		
RL	[2]		
RN	SEQUENCE FROM N.A.		
RP			

MEDLINE=21848401: PubMed-11859360:
 Wood V., Guilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sguoraks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Holtroyd S., Hornsby I., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 Weltens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 Coffeau A., Cadiou E., Dreano S., Gloux S., Lelaure V., Mortier S.,
 Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.:
 "The genome sequence of *Schizosaccharomyces pombe*.",
 Nature 415:871-880(2002).
 [3]
 SUBUNIT.
 RC STRAIN-972;
 RX MEDLINE-99144318; PubMed-9990507;
 RA Komlnami K.-I., Ochotorena I., Toda T.:
 "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
 complexes together with cullin-1 in fission yeast SCF (Skip-
 cullin-1-F-box) ubiquitin ligase.",
 RT Genes Cells 3:721-735(1998).
 RL [4]
 FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
 RC MEDLINE-22325332; PubMed-12167473;
 RX Seibert V., Prohl C., Schultze I., Rhee E., Lopez R., Abderazak K.,
 RA Zhou C., Wolf D.A.:
 "Combinatorial diversity of fission yeast SCF ubiquitin ligases by
 RT homo- and heterooligomeric assemblies of the F-box proteins Pop1 and
 RT Pop2p.",
 RL BMC Biochem. 3:222(2002).
 CC !- FUNCTION: Involved in maintenance of ploidy through proteasome-
 dependent degradation of CDK inhibitor rum1 and S-phase initiator
 cdc18. Functions as a recognition factor for rum1 and cdc18, which
 CC are subsequently ubiquitinated and targeted to the 26S proteasome
 CC for degradation.
 CC !- SUBUNIT: Homodimer and heterodimer with pop2. Binds to pcu1, pip1
 CC and cdc18.
 CC !- SUBCELLULAR LOCATION: Nuclear.
 CC !- SIMILARITY: Contains 1 F-box domain.
 CC !- SIMILARITY: Contains 5 WD repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).

Query Match 775 AA; 87816 MW; B06EDBA46553EEC1 CRC64; 76.2%; Score 16; DB 1; Length 775;

Qy 1 HXXXXXXH 7

Db 214 HSSSSH 220

RESULT 13

ID VGLB_HCMVA STANDARD; PRT; 906 AA.
 AC P06473;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB OR UL55.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=87080291; PubMed=3024973;
 RA Cranage M.P., Kouzarides T., Bankier A.T., Satchwell S.C., Weston K.,
 RA Tomlinson P., Barrell B.G., Hart H., Bell S.E., Minson A.C.,
 RA Smith G.L.;
 RT Identification of the human cytomegalovirus glycoprotein B gene and
 RT induction of neutralizing antibodies via its expression in
 RT recombinant vaccinia virus.*;
 RL EMBO J. 5:3057-3063(1986).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=87151139; PubMed=3029980;
 RA Kouzarides T., Bankier A.T., Satchwell S.C., Weston K.,
 RA Tomlinson P., Barrell B.G.;
 RT Large-scale rearrangement of homologous regions in the genomes of
 RT HCMV and EBV.*;
 RL Virology 157:397-413(1987).
 [3]
 RN SEQUENCE OF 735-906 FROM N.A.
 RP MEDLINE=87061230; PubMed=3023690;
 RA Kouzarides T., Bankier A.T., Satchwell S.C., Weston K., Tomlinson P.,
 RA Barrell B.G.;
 RT sequence and transcription analysis of the human cytomegalovirus DNA
 RT polymerase gene.*;
 RL J. Virol. 61:125-133(1987).
 [4]
 RN COMPLETE GENOME.
 RP MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169.*;
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X17403; CAA35414.1; -
 CC EMBL: X04606; CAA28274.1; -
 CC EMBL: M17209; AAA46009.1; -
 CC EMBL: M14709; AAA45987.1; -
 CC PIR: A25365; VGBEC1.
 CC InterPro: IPR000234; Glycoprot_B.
 CC Pfam: PF00606; Glycoprotein_B; 1.
 CC ProDom: PD000693; Glycoprot_B; 1.
 CC Glycoprotein; Transmembrane; Signal.
 CC SIGNAL 1 24
 CC CHAIN 25 906 GLYCOPROTEIN B.
 CC DOMAIN 25 706 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 707 722 POTENTIAL.
 FT TRANSMEM 728 747 POTENTIAL.
 FT TRANSMEM 751 771 POTENTIAL.
 FT DOMAIN 772 906 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 383 383 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 452 452 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 554 554 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 585 585 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 906 AA; 102003 MW; D4C7A6A3C7083FEE CRC64;

Query Match 76.2%; Score 16; DB 1; Length 906;

Best Local Similarity 28.6%; Pred. No. 1.4e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXH 7

Db 30 HATSSSTH 36

RESULT 14

ID VGLB_HCMVT STANDARD; PRT; 907 AA.
 AC P13201;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor [Contains: Glycoprotein GP55].
 GN GB OR UL55.
 OS Human cytomegalovirus (strain Towne).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89045645; PubMed=2460994;
 RA Spaete R.K., Thayer R.M., Probert W.S., Maslarz F.R.,
 RA Chamberlain S.H., Rasmussen L., Merigan T.C., Pacht C.;
 RT "Human cytomegalovirus strain Towne glycoprotein B is processed by
 RT proteolytic cleavage.*";
 RL Virology 167:207-225(1988).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M22343; AAA45920.1; -
 CC PIR: A31288; VGBETE.
 CC InterPro: IPR000234; Glycoprot_B.
 CC Pfam: PF00606; Glycoprotein_B; 1.
 CC ProDom: PD000693; Glycoprot_B; 1.
 CC Glycoprotein; Transmembrane; Signal.
 CC SIGNAL 1 24
 CC CHAIN 25 907 GLYCOPROTEIN B.


```
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF106037; AAF07395.1; ALT_INIT.
DR EMBL; AY028806; AAK37777.1; ALT_INIT.
DR EMBL; AY028807; AAK37778.1; ALT_INIT.
DR EMBL; AF183569; AAF20384.1; ALT_INIT.
DR EMBL; AF222340; AAF34664.1; ALT_INIT.
DR EMBL; AB011097; BAA25451.2; ALT_INIT.
DR EMBL; BC030775; BAA30775.1; ALT_INIT.
DR MIM; 606832; -.
DR MEROPS; M01.018; -.
DR InterPro; IPR001930; Ala_peptase.
DR InterPro; IPR006025; Zn_Mtpeptdse.
DR Pfam; PF01433; Peptidase_M1; 1.
DR PRINIS; PR00756; ALADIPASE.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Protease; Metalloprotease; Aminopeptidase; Zinc; Signal;
KW Glycoprotein; Alternative splicing; Polymorphism.
FT SIGNAL 1 24
FT CHAIN 25 929 ADIPOCYTE-DERIVED LEUCINE AMINOPEPTIDASE.
FT METAL 341 341 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 342 342 BY SIMILARITY.
FT METAL 345 345 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 464 464 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 426 426 PROTON DONOR (POTENTIAL).
FT CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 889 889 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 928 929 RM -> HDPEADATG (in isoform 2).
FT VARIANT 115 115 /FTIG-VSP_005450.
FT VARIANT 115 115 R -> P (IN dbSNP:26653).
FT VARIANT 264 264 /FTIG-VAR_012779.
FT VARIANT 334 334 I -> M (IN dbSNP:26618).
FT VARIANT 334 334 /FTIG-VAR_012780.
FT VARIANT 337 337 G -> D (IN dbSNP:27895).
FT VARIANT 337 337 /FTIG-VAR_012781.
FT VARIANT 516 516 M -> V (IN dbSNP:2287987).
FT VARIANT 516 516 /FTIG-VAR_012782.
FT VARIANT 718 718 K -> R (IN dbSNP:27434).
FT VARIANT 718 718 /FTIG-VAR_012783.
FT VARIANT 502 502 Q -> E (IN dbSNP:27044).
FT VARIANT 502 502 /FTIG-VAR_012784.
FT CONFLICT 502 502 G -> R (IN REF. 3).
FT CONFLICT 563 563 D -> N (IN REF. 2, 5 AND 6).
FT CONFLICT 713 713 R -> Q (IN REF. 2, 5 AND 6).
SQ SEQUENCE 929 AA; 105846 MW; 75C6AD58D0D70D51 CRC64;

Query Match 76.2%; Score 16; DB 1; Length 929;
Best Local Similarity 28.6%; Pred.No. 1.4e-03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
Db 491 HSSSSSH 497

Search completed: October 1, 2003, 10:42:40
Job time : 10.1667 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:33:32 ; Search time 41.8056 Seconds
(without alignments)
43.209 Million cell updates/sec

Title: US-10-057-890A-5
Perfect score: 21
Sequence: 1 HXXXXXXH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	76.2	51	13 Q98TK7	Q98tk7 platicthys
2	16	76.2	64	12 Q86837	Q86837 human cytom
3	16	76.2	64	12 Q86836	Q86836 human cytom
4	16	76.2	71	12 Q56639	Q56639 human cytom
5	16	76.2	71	12 Q56642	Q56642 human cytom
6	16	76.2	77	12 Q56885	Q56885 human cytom
7	16	76.2	83	12 Q69178	Q69178 human cytom
8	16	76.2	83	12 Q69179	Q69179 human cytom
9	16	76.2	83	12 Q69182	Q69182 human cytom
10	16	76.2	83	12 Q69181	Q69181 human cytom
11	16	76.2	86	12 Q68779	Q68779 human cytom
12	16	76.2	87	12 Q68785	Q68785 human cytom
13	16	76.2	87	12 Q68786	Q68786 human cytom
14	16	76.2	87	12 Q68782	Q68782 human cytom
15	16	76.2	87	12 Q68784	Q68784 human cytom
16	16	76.2	87	12 Q68787	Q68787 human cytom

17	16	76.2	87	12 Q68783	Q68783 human cytom
18	16	76.2	87	12 Q68774	Q68774 human cytom
19	16	76.2	87	12 Q68776	Q68776 human cytom
20	16	76.2	87	12 Q68781	Q68781 human cytom
21	16	76.2	87	12 Q68775	Q68775 human cytom
22	16	76.2	87	12 Q68773	Q68773 human cytom
23	16	76.2	87	12 Q68777	Q68777 human cytom
24	16	76.2	87	12 Q68778	Q68778 human cytom
25	16	76.2	87	12 Q68780	Q68780 human cytom
26	16	76.2	87	12 Q8YF67	Q8Yf67 brucella me
27	16	76.2	94	16 Q8G2Q5	Q8g2q5 brucella su
28	16	76.2	115	4 Q9P1D1	Q9p1d1 homo sapien
29	16	76.2	115	12 Q9YK12	Q9Yk12 human cytom
30	16	76.2	115	12 Q9YKH9	Q9Ykh9 human cytom
31	16	76.2	115	12 Q9WIL1	Q9Wll1 human cytom
32	16	76.2	115	12 Q9YK11	Q9Yk11 human cytom
33	16	76.2	115	12 Q9YKH6	Q9Ykh6 human cytom
34	16	76.2	115	12 Q9W8H0	Q9W8h0 human cytom
35	16	76.2	115	12 Q9WA31	Q9wa31 human cytom
36	16	76.2	115	12 Q9YKH8	Q9Ykh8 human cytom
37	16	76.2	115	12 Q9W9G7	Q9W9g7 human cytom
38	16	76.2	115	12 Q9YKH7	Q9Ykh7 human cytom
39	16	76.2	115	12 Q9YK10	Q9Yk10 human cytom
40	16	76.2	115	12 Q9YKH5	Q9Ykh5 human cytom
41	16	76.2	115	12 Q9WA39	Q9wa39 human cytom
42	16	76.2	115	12 Q9WIL0	Q9Wll0 human cytom
43	16	76.2	115	12 Q9W9K7	Q9W9k7 human cytom
44	16	76.2	116	12 Q9YKH1	Q9Ykh1 human cytom
45	16	76.2	116	12 Q9YKH2	Q9Ykh2 human cytom

ALIGNMENTS

RESULT 1
Q98TK7 PRELIMINARY: PRT: 51 AA.
AC Q98TK7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Putative mucin (fragment).
MDC.
GN Platicthys flesus (European flounder).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Platicthys.
OX NCBI_TaxID=8260;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Williams T.D., Chipman J.K.;
RT "A DNA array to monitor the effects of environmental pollution on European flounder (Platicthys flesus).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
FR EMBL; AJ306238; CAC27527.1; .
FT NON_TER 1 1
FT NON_TER 51 51
SQ SEQUENCE 51 AA; 4982 MW; F6D750E09A15C24B CRC64;
Query Match 76.2%; Score 16; DB 13; Length 51;
Best Local similarity 28.6%; Pred. No. 1.1e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 HXXXXXXH 7
Db 39 HSSSSSTH 45
RESULT 2
Q86837 PRELIMINARY: PRT: 64 AA.
ID Q86837

AC Q86837;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Glycoprotein B (Fragment).
 GN GB.
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95053935; PubMed=7525866;
 RA Ayata M., Sugano T., Murayama T., Sakamuro D., Takegami T.,
 RA Matsumoto Y., Furukawa T.;
 RT "Different antibody response to a neutralizing epitope of human
 RT cytomegalovirus glycoprotein B among seropositive individuals.";
 RL J. Med. Virol. 43:386-392(1994).
 DR EMBL; S74441; AAB32493.2; -;
 FT NON_TER 1 1
 FT NON_TER 64 64
 SQ SEQUENCE 64 AA; 6874 MW; 99483C5518CC5199 CRC64;

Query Match 76.2%; Score 16; DB 12; Length 64;
 Best Local Similarity 28.6%; Pred. No. 1.3e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXXH 7
 | |
 Db 5 HATSSTH 11

RESULT 3
 Q86836 PRELIMINARY; PRT; 64 AA.
 AC Q86836;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Glycoprotein B (Fragment).
 GN GB.
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95053935; PubMed=7525866;
 RA Ayata M., Sugano T., Murayama T., Sakamuro D., Takegami T.,
 RA Matsumoto Y., Furukawa T.;
 RT "Different antibody response to a neutralizing epitope of human
 RT cytomegalovirus glycoprotein B among seropositive individuals.";
 RL J. Med. Virol. 43:386-392(1994).
 DR EMBL; S74440; AAB32492.2; -;
 FT NON_TER 1 1
 FT NON_TER 64 64
 SQ SEQUENCE 64 AA; 6784 MW; 9941452518CC5E3C CRC64;

Query Match 76.2%; Score 16; DB 12; Length 64;
 Best Local Similarity 28.6%; Pred. No. 1.3e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXXH 7
 | |
 Db 5 HATSSAH 11

RESULT 4
 O56639 PRELIMINARY; PRT; 71 AA.
 AC O56639;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Glycoprotein B (Fragment).
 DE Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=918-91;
 RA Shepp D.H., Match M.E., Lipson S.M., Pergolizzi R.G.;
 RT "A fifth human cytomegalovirus glycoprotein B genotype.";
 RL Res. Virol. 0:0-0(1998).
 DR EMBL; AF043073; AAB99833.1; -;
 FT NON_TER 1 1
 FT NON_TER 71 71
 SQ SEQUENCE 71 AA; 7472 MW; BCE457D2D181B6A2 CRC64;

Query Match 76.2%; Score 16; DB 12; Length 71;
 Best Local Similarity 28.6%; Pred. No. 1.4e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXXH 7
 | |
 Db 29 HATSSAH 35

RESULT 5
 O56642 PRELIMINARY; PRT; 71 AA.
 AC O56642;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Glycoprotein B (Fragment).
 DE Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=219-91;
 RA Shepp D.H., Match M.E., Lipson S.M., Pergolizzi R.G.;
 RT "A fifth human cytomegalovirus glycoprotein B genotype.";
 RL Res. Virol. 0:0-0(1998).
 DR EMBL; AF043076; AAB99836.1; -;
 FT NON_TER 1 1
 FT NON_TER 71 71
 SQ SEQUENCE 71 AA; 7449 MW; BCE457D2D421B6A2 CRC64;

Query Match 76.2%; Score 16; DB 12; Length 71;
 Best Local Similarity 28.6%; Pred. No. 1.4e+03;
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXXXH 7
 | |
 Db 29 HATSSAH 35

RESULT 6
 O56885 PRELIMINARY; PRT; 77 AA.
 AC O56885;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Glycoprotein B (Fragment).
 GN GB.
 OS Human cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10359;
 RN [1]
 RP SEQUENCE FROM N.A.

```
RC STRAIN-P8/LUNG:
RT Meyer-Knig U., Ebert K., Schrage B., Pollak S., Hufert F.T.;
RA "Different Human cytomegalovirus strains can establish latency in a
RL single healthy subject.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y16543; CAA76281.1; -.
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8430 MW; 65E2F6B34D345DA7 CRC64;

Query Match 76.2%; Score 16; DB 12; Length 77;
Best Local Similarity 28.6%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
DB 8 HTTSAAH 14

RESULT 7
Q69178 PRELIMINARY; PRT; 83 AA.
AC Q69178;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Glycoprotein gp58/116 (Fragment).
GN UL55.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92129518; PubMed-1663509;
RA Lehner R., Stamminger T., Mach M.;
RT "Comparative sequence analysis of human cytomegalovirus strains.";
RL J. Clin. Microbiol. 29:2494-2502(1991).
DR EMBL: M64939; AAA45943.1; -.
FT NON_TER 83
FT NON_TER 83
SQ SEQUENCE 83 AA; 8757 MW; EB5582A7ED81377F CRC64;

Query Match 76.2%; Score 16; DB 12; Length 83;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
DB 30 HATSSAH 36

RESULT 8
Q69179 PRELIMINARY; PRT; 83 AA.
AC Q69179;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Glycoprotein gp58/116 (Fragment).
GN UL55.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92129518; PubMed-1663509;
RA Lehner R., Stamminger T., Mach M.;
RT "Comparative sequence analysis of human cytomegalovirus strains.";
RL J. Clin. Microbiol. 29:2494-2502(1991).

RC STRAIN-P8/LUNG:
RT Meyer-Knig U., Ebert K., Schrage B., Pollak S., Hufert F.T.;
RA "Different Human cytomegalovirus strains can establish latency in a
RL single healthy subject.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y16543; CAA76281.1; -.
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8430 MW; 65E2F6B34D345DA7 CRC64;

Query Match 76.2%; Score 16; DB 12; Length 77;
Best Local Similarity 28.6%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
DB 8 HTTSAAH 14

RESULT 7
Q69178 PRELIMINARY; PRT; 83 AA.
AC Q69178;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Glycoprotein gp58/116 (Fragment).
GN UL55.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92129518; PubMed-1663509;
RA Lehner R., Stamminger T., Mach M.;
RT "Comparative sequence analysis of human cytomegalovirus strains.";
RL J. Clin. Microbiol. 29:2494-2502(1991).
DR EMBL: M64939; AAA45943.1; -.
FT NON_TER 83
FT NON_TER 83
SQ SEQUENCE 83 AA; 8757 MW; EB5582A7ED81377F CRC64;

Query Match 76.2%; Score 16; DB 12; Length 83;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
DB 30 HATSSAH 36

RESULT 8
Q69179 PRELIMINARY; PRT; 83 AA.
AC Q69179;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Glycoprotein gp58/116 (Fragment).
GN UL55.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92129518; PubMed-1663509;
RA Lehner R., Stamminger T., Mach M.;
RT "Comparative sequence analysis of human cytomegalovirus strains.";
RL J. Clin. Microbiol. 29:2494-2502(1991).

DR EMBL: M64937; AAA45941.1; -.
FT NON_TER 83
SQ SEQUENCE 83 AA; 8694 MW; E083FBB72DAA1503 CRC64;

Query Match 76.2%; Score 16; DB 12; Length 83;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
DB 41 HTTSAAH 47

RESULT 9
Q69182 PRELIMINARY; PRT; 83 AA.
AC Q69182;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Glycoprotein gp58/116 (Fragment).
GN UL55.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92129518; PubMed-1663509;
RA Lehner R., Stamminger T., Mach M.;
RT "Comparative sequence analysis of human cytomegalovirus strains.";
RL J. Clin. Microbiol. 29:2494-2502(1991).
DR EMBL: M64956; AAA45945.1; -.
FT NON_TER 83
FT NON_TER 83
SQ SEQUENCE 83 AA; 8858 MW; F45A92BD5D87207F CRC64;

Query Match 76.2%; Score 16; DB 12; Length 83;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
DB 30 HATSSH 36

RESULT 10
Q69181 PRELIMINARY; PRT; 83 AA.
AC Q69181;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Glycoprotein gp58/116 (Fragment).
GN UL55.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92129518; PubMed-1663509;
RA Lehner R., Stamminger T., Mach M.;
RT "Comparative sequence analysis of human cytomegalovirus strains.";
RL J. Clin. Microbiol. 29:2494-2502(1991).
DR EMBL: M64949; AAA45944.1; -.
FT NON_TER 83
FT NON_TER 83
SQ SEQUENCE 83 AA; 8694 MW; E083FBB72DAA1503 CRC64;

Query Match 76.2%; Score 16; DB 12; Length 83;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
```

```

Db      41 HTTSAAH 47

RESULT 11
Q68779          PRELIMINARY;      PRT;      86 AA.
ID O68779;
AC O68779;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE Glycoprotein B, amino part of (Fragment).
GN GB, 5' REGION.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95070784; PubMed=7979986;
RA Shiu S.Y., Chan K.M., Lo S.K., Ip K.W., Yuen K.Y., Heath R.B.;
RT "Sequence variation of the amino-terminal antigenic domains of
RT glycoprotein B of human cytomegalovirus strains isolated from Chinese
RT patients.;"
RL Arch. Virol. 137:133-138(1994).
DR EMBL; X76152; CAA53750.1; -.
FT NON_TER 1
FT NON_TER 86
FT SEQUENCE 86 AA; 9091 MW; 8113880715880C69 CRC64;
SQ

Query Match 76.2%; Score 16; DB 12; Length 86;
Best Local Similarity 28.6%; Pred. NO. 1.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
Db 23 HATSTTH 29

RESULT 12
Q68785          PRELIMINARY;      PRT;      87 AA.
ID O68785;
AC O68785;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE Glycoprotein B, amino part of (Fragment).
GN GB, 5' REGION.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95070784; PubMed=7979986;
RA Shiu S.Y., Chan K.M., Lo S.K., Ip K.W., Yuen K.Y., Heath R.B.;
RT "Sequence variation of the amino-terminal antigenic domains of
RT glycoprotein B of human cytomegalovirus strains isolated from Chinese
RT patients.;"
RL Arch. Virol. 137:133-138(1994).
DR EMBL; X76146; CAA53744.1; -.
FT NON_TER 1
FT NON_TER 87
FT SEQUENCE 87 AA; 9178 MW; EA91C9DDFF3D5C0C CRC64;
SQ

Query Match 76.2%; Score 16; DB 12; Length 87;
Best Local Similarity 28.6%; Pred. NO. 1.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
Db 24 HATSTTH 30

```

```

RESULT 13
Q68786          PRELIMINARY;      PRT;      87 AA.
ID O68786;
AC O68786;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE Glycoprotein B, amino part of (Fragment).
GN GB, 5' REGION.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95070784; PubMed=7979986;
RA Shiu S.Y., Chan K.M., Lo S.K., Ip K.W., Yuen K.Y., Heath R.B.;
RT "Sequence variation of the amino-terminal antigenic domains of
RT glycoprotein B of human cytomegalovirus strains isolated from Chinese
RT patients.;"
RL Arch. Virol. 137:133-138(1994).
DR EMBL; X76155; CAA53753.1; -.
FT NON_TER 1
FT NON_TER 87
FT SEQUENCE 87 AA; 9097 MW; CF8E50779917E86B CRC64;
SQ

Query Match 76.2%; Score 16; DB 12; Length 87;
Best Local Similarity 28.6%; Pred. NO. 1.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
Db 35 HTTSAAH 41

RESULT 14
Q68782          PRELIMINARY;      PRT;      87 AA.
ID O68782;
AC O68782;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE Glycoprotein B, amino part of (Fragment).
GN GB, 5' REGION.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95070784; PubMed=7979986;
RA Shiu S.Y., Chan K.M., Lo S.K., Ip K.W., Yuen K.Y., Heath R.B.;
RT "Sequence variation of the amino-terminal antigenic domains of
RT glycoprotein B of human cytomegalovirus strains isolated from Chinese
RT patients.;"
RL Arch. Virol. 137:133-138(1994).
DR EMBL; X76153; CAA53751.1; -.
FT NON_TER 1
FT NON_TER 87
FT SEQUENCE 87 AA; 9113 MW; BE524DAB9903FD2A CRC64;
SQ

Query Match 76.2%; Score 16; DB 12; Length 87;
Best Local Similarity 28.6%; Pred. NO. 1.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXX 7
|
Db 35 HTTSAAH 41

RESULT 15
Q68784

```


ID Q68784 PRELIMINARY; PRT; 87 AA.
AC Q68784;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Glycoprotein B, amino part of (Fragment).
GN GB, 5' REGION.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95070784; PubMed-7979986;
RA Shiu S.Y., Chan K.M., Lo S.K., Ip K.W., Yuen K.Y., Heath R.B.;
RT "Sequence variation of the amino-terminal antigenic domains of
RT glycoprotein B of human cytomegalovirus strains isolated from Chinese
RL Arch. Virol. 137:133-138(1994).
RL EMBL; X76145; CAA53743.1; -.
FT NON_TER 1 87
FT NON_TER 87 87
SQ SEQUENCE 87 AA: 9277 MW; 1F91DC8FF3D5C16 CRC64;

Query Match 76.2%; Score 16; DB 12; Length 87;
Best Local Similarity 28.6%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Caps 0;

QY 1 HXXXXXH 7
Db 24 HATSSTH 30

Search completed: October 1, 2003, 10:46:26
Job time : 43.8056 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:38:06 ; Search time 14.5833 Seconds
(without alignments)
20.309 Million cell updates/sec

Title: US-10-057-890A-5
Perfect score: 21
Sequence: 1 HXXXXXXH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	76.2	135	4	US-09-199-637A-217
2	16	76.2	629	4	US-09-252-991A-27100
3	16	76.2	906	1	US-08-220-151-9
4	16	76.2	906	1	US-08-413-118-9
5	16	76.2	906	3	US-08-473-446-9
6	16	76.2	907	1	US-08-349-006-2
7	16	76.2	907	3	US-08-804-439A-19
8	16	76.2	907	3	US-08-720-229-19
9	16	76.2	907	4	US-09-171-699-2
10	16	76.2	907	5	PCT-US94-04180-2
11	16	76.2	941	4	US-09-996-243-353
12	16	76.2	948	4	US-09-620-312D-1105
13	16	76.2	1357	4	US-09-252-991A-21308
14	15	71.4	7	3	US-08-478-316-49
15	15	71.4	7	4	US-09-019-793A-49
16	15	71.4	15	3	US-09-153-586-26
17	15	71.4	16	4	US-09-006-755B-7
18	15	71.4	17	1	US-08-305-871A-7
19	15	71.4	17	4	US-08-788-822A-8
20	15	71.4	17	5	PCT-US94-09795-3
21	15	71.4	18	3	US-08-810-009-33
22	15	71.4	18	3	US-08-810-009-34
23	15	71.4	18	3	US-08-810-009-37
24	15	71.4	18	3	US-08-810-009-38
25	15	71.4	18	3	US-08-810-009-40
26	15	71.4	27	3	US-09-153-586-6
27	15	71.4	82	4	US-08-311-731A-73

28	15	71.4	104	2	US-08-727-688-32	Sequence 32, Appl
29	15	71.4	111	4	US-09-252-991A-26598	Sequence 26598, A
30	15	71.4	130	4	US-09-464-535-20	Sequence 20, Appl
31	15	71.4	134	4	US-08-311-731A-196	Sequence 196, App
32	15	71.4	136	4	US-09-252-991A-18247	Sequence 18247, A
33	15	71.4	137	1	US-07-953-230A-10	Sequence 10, Appl
34	15	71.4	143	4	US-09-439-313-478	Sequence 478, App
35	15	71.4	153	1	US-08-219-878A-1	Sequence 1, Appl
36	15	71.4	153	4	US-09-252-991A-25216	Sequence 25216, A
37	15	71.4	153	5	PCT-US93-04329-1	Sequence 1, Appl
38	15	71.4	154	4	US-09-252-991A-19388	Sequence 19388, A
39	15	71.4	156	3	US-09-142-583A-11	Sequence 11, Appl
40	15	71.4	168	6	5194425-4	Patent No. 5194425
41	15	71.4	170	1	US-08-227-372-1	Sequence 1, Appl
42	15	71.4	170	6	5194425-3	Patent No. 5194425
43	15	71.4	175	4	US-09-252-991A-29157	Sequence 29157, A
44	15	71.4	188	3	US-08-470-335-204	Sequence 204, App
45	15	71.4	190	3	US-08-341-018-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-09-199-637A-217
; Sequence 217, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Rui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULANCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 217
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-217
Query Match 76.2% Score 16; DB 4; Length 135;
Best Local Similarity 28.6%; Pred. No. 2.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 HXXXXXXH 7
| | |
Db 7 HSSAAAH 13
RESULT 2
US-09-252-991A-27100
; Sequence 27100, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMON
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27100
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27100

Query Match 76.2%; Score 16; DB 4; Length 629;
Best Local Similarity 28.6%; Pred. No. 6.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXX 7
|
Db 128 HTATASH 134

RESULT 3
US-08-220-151-9
; Sequence 9, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Limbach, Keith J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-220-151-9

Query Match 76.2%; Score 16; DB 1; Length 906;
Best Local Similarity 28.6%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXX 7
|
Db 30 HATSSTH 16

RESULT 4
US-08-413-118-9

; Sequence 9, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-413-118-9

Query Match 76.2%; Score 16; DB 1; Length 906;
Best Local Similarity 28.6%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXX 7
|
Db 30 HATSSTH 36

RESULT 5
US-08-473-446-9
; Sequence 9, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC, AND gD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
;
US-08-473-446-9

Query Match 76.2%; Score 16; DB 3; Length 906;
Best Local Similarity 28.6%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
DB 30 HATSSTH 36

RESULT 6
US-08-349-006-2
; Sequence 2, Application US/08349006
; Patent No. 5552143
; GENERAL INFORMATION:
; APPLICANT: Plotkin, Stanley A.
; APPLICANT: Ricciardi, Robert P.
; APPLICANT: Gonczol, Eva
; APPLICANT: Berencsi, Klara
; APPLICANT: Rando, Robert F.
; TITLE OF INVENTION: Recombinant Cytomegalovirus Vaccine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,006
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/048,978
; FILING DATE: 19930416
; APPLICATION NUMBER: US 07/328,406
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/666,471
; FILING DATE: 06-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/017,130

```

```

; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST6C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-349-006-2

Query Match 76.2%; Score 16; DB 1; Length 907;
Best Local Similarity 28.6%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
DB 41 HTTSAH 47

RESULT 7
US-08-804-439A-19
; Sequence 19, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TEXT:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-804-439A-19

Query Match 76.2%; Score 16; DB 3; Length 907;
Best Local Similarity 28.6%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy 1 HXXXXX 7
|
Db 41 HTTSAAH 47

RESULT 8
US-08-720-229-19
; Sequence 19, Application US/08720229
; Patent No. 6022542
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Kurt
; APPLICANT: Strand, Mark L.
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,229
; FILING DATE: 26-SEP-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-720-229-19

Query Match 76.2%; Score 16; DB 3; Length 907;
Best Local Similarity 28.6%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXX 7
|
Db 41 HTTSAAH 47

RESULT 9
US-09-171-699-2
; Sequence 2, Application US/09171699
; Patent No. 6448389
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Gonczol, Eva
; Berencsi, Klara
; Kari, Csaba
; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and
; Uses Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457

; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,699
; FILING DATE: 19-Jan-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/015,717
; FILING DATE: 23-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST66APCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-699-2

Query Match 76.2%; Score 16; DB 4; Length 907;
Best Local Similarity 28.6%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXX 7
|
Db 41 HTTSAAH 47

RESULT 10
PCT-US94-04180-2
; Sequence 2, Application PC/TUS9404180
; GENERAL INFORMATION:
; APPLICANT: Wistar Institute of Anatomy, Biology
; APPLICANT: Government of USA, Dept.
; APPLICANT: Health and Human, Services
; TITLE OF INVENTION: Recombinant Cytomegalovirus Vaccine
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04180
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,978
; FILING DATE: 16-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST6CPCT

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-04180-2

Query Match 76.2%; Score 16; DB 5; Length 907;
Best Local Similarity 28.6%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HXXXXH 7
Db 41 HTTSAAH 47

RESULT 11
US-09-996-243-353
; Sequence 353, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geurtsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
```

; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982

; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 76.2%; Score 16; DB 4; Length 941;
Best Local Similarity 28.6%; Pred. No. 8.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXX 7
DB 503 HSSSSS 509

RESULT 12

US-09-620-312D-1105
; Sequence 1105, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 1105
; LENGTH: 948
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-620-312D-1105

Query Match 76.2%; Score 16; DB 4; Length 948;
Best Local Similarity 28.6%; Pred. No. 8.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXX 7
DB 503 HSSSSS 509

RESULT 13

US-09-252-991A-21308
; Sequence 21308, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21308
; LENGTH: 1357
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21308

Query Match          76.2%; Score 16; DB 4; Length 1357;
Best Local Similarity 28.6%; Pred. No. 1.1e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HXXXXXH 7
Db      634 HTATASH 640

RESULT 14
US-08-478-316-49
; Sequence 49, Application US/08478316
; Patent No. 6251397
; GENERAL INFORMATION:
; APPLICANT: PAUL, PREM S.
; APPLICANT: HALBUR, PATRICK G.
; APPLICANT: MENG, XIANG-JIN
; APPLICANT: MOROZOV, IGOR
; TITLE OF INVENTION: POLYNUCLEIC ACIDS ISOLATED FROM A PORCINE
; TITLE OF INVENTION: REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS (PRRSV), PROTEINS
; TITLE OF INVENTION: ENCODED BY THE POLYNUCLEIC ACIDS, VACCINES BASED ON THE
; TITLE OF INVENTION: PROTEINS AND/OR POLYNUCLEIC ACIDS, A METHOD OF PROTECTING A
; TITLE OF INVENTION: PIG FROM PRRS AND A METHOD OF DETECTING A PRRSV
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,316
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/301,435
; FILING DATE: 01-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/131,625
; FILING DATE: 05-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/969,071
; FILING DATE: 30-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M.P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 4625-026-55X CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS:
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-478-316-49

Query Match          71.4%; Score 15; DB 3; Length 7;
Best Local Similarity 28.6%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HXXXXXH 7
Db      1 HSNSSSH 7

RESULT 15
US-09-019-793A-49
; Sequence 49, Application US/09019793A
; Patent No. 6380376
; GENERAL INFORMATION:
; APPLICANT: PAUL, Prem
; APPLICANT: MENG, Xiang-Jin
; APPLICANT: MOROZOV, Igor
; APPLICANT: HALBUR, Patrick
; TITLE OF INVENTION: PROTEINS ENCODED BY POLYNUCLEIC ACIDS OF PORCINE
; TITLE OF INVENTION: REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS (PRRSV)
; FILE REFERENCE: 4625-0039-55X CIP
; CURRENT APPLICATION NUMBER: US/09/019,793A
; CURRENT FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: 08/478,316
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/301,435
; PRIOR FILING DATE: 1994-09-01
; PRIOR APPLICATION NUMBER: 08/131,625
; PRIOR FILING DATE: 1993-10-05
; PRIOR APPLICATION NUMBER: 07/969,071
; PRIOR FILING DATE: 1992-10-30
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Peptide
US-09-019-793A-49

Query Match          71.4%; Score 15; DB 4; Length 7;
Best Local Similarity 28.6%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HXXXXXH 7
Db      1 HSNSSSH 7

Search completed: October 1, 2003, 10:49:11
Job time : 15.5833 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:42:07 ; Search time 18.6667 seconds
(without alignments)
59.330 Million cell updates/sec

Title: US-10-057-890A-5
Perfect score: 21
Sequence: 1 HXXXXXXH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	76.2	135	11	US-09-975-719-217
2	16	76.2	170	11	US-09-934-455-124
3	16	76.2	290	9	US-09-764-853-661
4	16	76.2	307	12	US-10-287-274-362
5	16	76.2	514	10	US-09-738-626-5869
6	16	76.2	596	15	US-10-205-823-222
7	16	76.2	764	15	US-10-081-119-30
8	16	76.2	907	15	US-10-223-538-2
9	16	76.2	930	11	US-09-866-050A-636
10	16	76.2	930	11	US-09-866-050A-644
11	16	76.2	941	9	US-09-726-643-60
12	16	76.2	941	9	US-09-989-722-353
13	16	76.2	941	9	US-09-989-723-353
14	16	76.2	941	9	US-09-989-729-353
15	16	76.2	941	9	US-09-989-727-353

16	76.2	941	10	US-09-989-731-353	Sequence 353, App
17	76.2	941	10	US-09-989-732-353	Sequence 353, App
18	76.2	941	10	US-09-991-073-353	Sequence 353, App
19	76.2	941	10	US-09-990-442-353	Sequence 353, App
20	76.2	941	10	US-09-991-163-353	Sequence 353, App
21	76.2	941	10	US-09-993-604-353	Sequence 353, App
22	76.2	941	10	US-09-990-456-353	Sequence 353, App
23	76.2	941	10	US-09-989-721-353	Sequence 353, App
24	76.2	941	10	US-09-992-598-353	Sequence 353, App
25	76.2	941	10	US-09-988-293A-353	Sequence 353, App
26	76.2	941	10	US-09-989-735-353	Sequence 353, App
27	76.2	941	10	US-09-990-444-353	Sequence 353, App
28	76.2	941	10	US-09-991-181-353	Sequence 353, App
29	76.2	941	10	US-09-989-730-353	Sequence 353, App
30	76.2	941	10	US-09-990-436-353	Sequence 353, App
31	76.2	941	10	US-09-993-687-353	Sequence 353, App
32	76.2	941	11	US-09-989-734-353	Sequence 353, App
33	76.2	941	11	US-09-997-653-353	Sequence 353, App
34	76.2	941	11	US-09-993-667-353	Sequence 353, App
35	76.2	941	11	US-09-997-428-353	Sequence 353, App
36	76.2	941	11	US-09-997-666-353	Sequence 353, App
37	76.2	941	11	US-09-990-438-353	Sequence 353, App
38	76.2	941	11	US-09-990-562-353	Sequence 353, App
39	76.2	941	11	US-09-990-711-353	Sequence 353, App
40	76.2	941	11	US-09-998-726-353	Sequence 353, App
41	76.2	941	11	US-09-998-156-353	Sequence 353, App
42	76.2	941	11	US-09-990-437-353	Sequence 353, App
43	76.2	941	11	US-09-991-157-353	Sequence 353, App
44	76.2	941	11	US-09-997-514-353	Sequence 353, App
45	76.2	941	11	US-09-997-573-353	Sequence 353, App

ALIGNMENTS

RESULT 1
US-09-975-719-217
; Sequence 217, Application US/09975719
; Publication No. US20030022349A1
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Rahme, Laurence G.
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361003
; CURRENT APPLICATION NUMBER: US/09/975,719
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 09/199,637
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: US 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 217
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-975-719-217

Query Match 76.2% Score 16; DB 11; Length 135;
Best Local Similarity 28.6%; Pred. No. 5.3e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXXH 7
| | |
Db 7 HSSAAAH 13

RESULT 2
US-09-934-455-124
; Sequence 124, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:

```

; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omalra
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; CURRENT FILING DATE: 2001-08-22
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 124
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-124

```

```

Query Match          76.2%; Score 16; DB 11; Length 170;
Best Local Similarity 28.6%; Pred. No. 6.1e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 1 HXXXXXX 7
|
Db 143 HAAAAAH 149

```

```

RESULT 3
US-09-764-853-661
; Sequence 661, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ06
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 661
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-661

```

```

Query Match          76.2%; Score 16; DB 9; Length 290;
Best Local Similarity 28.6%; Pred. No. 8.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 1 HXXXXXX 7
|
Db 145 HSSSSSH 151

```

```

RESULT 4
US-10-287-274-362
; Sequence 362, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn

```

```

; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THE
; FILE REFERENCE: ELITRA.008DVI
; CURRENT APPLICATION NUMBER: US/10/287,274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 362
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-362

```

```

Query Match          76.2%; Score 16; DB 12; Length 307;
Best Local Similarity 28.6%; Pred. No. 8.8e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 1 HXXXXXX 7
|
Db 70 HSSTASH 76

```

```

RESULT 5
US-09-738-626-5869
; Sequence 5869, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5869
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5869

```

```

Query Match          76.2%; Score 16; DB 10; Length 514;
Best Local Similarity 28.6%; Pred. No. 1.2e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 1 HXXXXXX 7
|
Db 343 HSAASAH 349

```

```

RESULT 6
US-10-205-823-222
; Sequence 222, Application US/10205823
; Publication No. US20030108963A1

```



```
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 636
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-636

Query Match          76.2%  Score 16; DB 11; Length 930;
Best Local Similarity 28.6%  Pred. No. 1.7e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXH 7
Db 492 HSSSTSH 498

RESULT 10
US-09-866-050A-644
; Sequence 644, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 644
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-644

Query Match          76.2%  Score 16; DB 11; Length 930;
Best Local Similarity 28.6%  Pred. No. 1.7e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXH 7
Db 492 HSSSTSH 498

RESULT 11
US-09-726-643-60
; Sequence 60, Application US/09726643
; Patent No. US20020028449A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040Pl
; CURRENT APPLICATION NUMBER: US/09/726,643
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 941
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (807)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (809)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (815)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (819)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-726-643-60

Query Match          76.2%  Score 16; DB 9; Length 941;
Best Local Similarity 28.6%  Pred. No. 1.8e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXXH 7
Db 503 HSSSSH 509

RESULT 12
US-09-989-722-353
; Sequence 353, Application US/09989722
; Patent No. US20020072067A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC63
; CURRENT APPLICATION NUMBER: US/09/989,722
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
```

7	PRIOR APPLICATION NUMBER: 60/089599	
7	PRIOR FILING DATE: 1998-06-17	
7	PRIOR APPLICATION NUMBER: 60/089600	
7	PRIOR FILING DATE: 1998-06-17	
7	PRIOR APPLICATION NUMBER: 60/089653	
7	PRIOR FILING DATE: 1998-06-17	
7	PRIOR APPLICATION NUMBER: 60/089801	
7	PRIOR FILING DATE: 1998-06-18	
7	PRIOR APPLICATION NUMBER: 60/089907	
7	PRIOR FILING DATE: 1998-06-18	
7	PRIOR APPLICATION NUMBER: 60/089908	
7	PRIOR FILING DATE: 1998-06-18	
7	PRIOR APPLICATION NUMBER: 60/089947	
7	PRIOR FILING DATE: 1998-06-19	
7	PRIOR APPLICATION NUMBER: 60/089948	
7	PRIOR FILING DATE: 1998-06-19	
7	PRIOR APPLICATION NUMBER: 60/089952	
7	PRIOR FILING DATE: 1998-06-19	
7	PRIOR APPLICATION NUMBER: 60/090246	
7	PRIOR FILING DATE: 1998-06-22	
7	PRIOR APPLICATION NUMBER: 60/090252	
7	PRIOR FILING DATE: 1998-06-22	
7	PRIOR APPLICATION NUMBER: 60/090254	
7	PRIOR FILING DATE: 1998-06-22	
7	PRIOR APPLICATION NUMBER: 60/090349	
7	PRIOR FILING DATE: 1998-06-23	
7	PRIOR APPLICATION NUMBER: 60/090355	
7	PRIOR FILING DATE: 1998-06-23	
7	PRIOR APPLICATION NUMBER: 60/090429	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090431	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090435	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090444	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090445	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090472	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090535	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090540	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090542	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090557	
7	PRIOR FILING DATE: 1998-06-24	
7	PRIOR APPLICATION NUMBER: 60/090676	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090678	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090690	
7	PRIOR FILING DATE: 1998-06-25	
7	PRIOR APPLICATION NUMBER: 60/090862	
7	PRIOR FILING DATE: 1998-06-26	
7	PRIOR APPLICATION NUMBER: 60/090863	
7	PRIOR FILING DATE: 1998-06-26	
7	PRIOR APPLICATION NUMBER: 60/091360	
7	PRIOR FILING DATE: 1998-07-01	
7	PRIOR APPLICATION NUMBER: 60/091478	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091544	
7	PRIOR FILING DATE: 1998-07-01	
7	PRIOR APPLICATION NUMBER: 60/091519	
7	PRIOR FILING DATE: 1998-07-02	
7	PRIOR APPLICATION NUMBER: 60/091626	

```

; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          76.2%  Score 16; DB 9; Length 941;
Best Local Similarity 28.6%  Pred.No. 1.8e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 HXXXXXX 7
Db      503 HSSSSSH 509

RESULT 13
US-09-989-723-353
; Sequence 353, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC62
; CURRENT APPLICATION NUMBER: US/09/989,723
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801

```

; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 76.2%; Score 16; DB 9; Length 941;
Best Local Similarity 28.6%; Pred. No. 1.8e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Cy 1 HXXXXX 7
Db 503 HSSSSSH 509

RESULT 14

US-09-989-279-353
; Sequence 353, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P27301C56
; CURRENT APPLICATION NUMBER: US/09/989,279
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827

;
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19

;
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090578
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 76.2%; Score 16; DB 9; Length 941;
Best Local Similarity 28.6%; Pred. No. 1.8e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 HXXXXX 7
Db 503 HSSSSH 509

RESULT 15
US-09-989-727-353
; Sequence 353, Application US/09989727
; Patent No. US20020072497A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC65
; CURRENT APPLICATION NUMBER: US/09/989,727
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252

; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 76.2%; Score 16; DB 9; Length 941;
Best Local Similarity 28.6%; Pred. No. 1.8e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Caps 0;

QY 1 HXXXXXX 7
DB 503 HSSSSH 509

Search completed: October 1, 2003, 10:50:53
Job time : 19.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:42:07 ; Search time 16 Seconds
(without alignments)
59,330 Million cell updates/sec

Title: US-10-057-890A-6

Perfect score: 40

Sequence: 1 YKGLC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	6	15	US-10-057-890A-6
2	40	100.0	138	15	US-10-057-890A-10
3	40	100.0	157	15	US-10-057-890A-31
4	36	90.0	3067	12	US-09-949-029-18
5	35	87.5	250	9	US-09-867-550-760
6	35	87.5	313	9	US-09-800-729-196
7	35	87.5	316	10	US-09-764-864-967
8	34	85.0	23	9	US-09-785-632A-39
9	34	85.0	23	12	US-10-223-765-39
10	34	85.0	308	10	US-09-764-864-927
11	34	85.0	309	10	US-09-764-864-1366
12	34	85.0	366	12	US-09-949-029-96
13	34	85.0	530	11	US-09-372-348-11
14	34	85.0	600	10	US-09-764-864-957
15	34	85.0	1323	10	US-09-801-368-34

16	33	82.5	55	10	US-09-796-892-2505	Sequence 2505, Ap
17	33	82.5	55	15	US-10-040-862-2505	Sequence 2505, Ap
18	33	82.5	129	10	US-09-764-864-1352	Sequence 1352, Ap
19	33	82.5	334	10	US-09-764-864-908	Sequence 908, Ap
20	33	82.5	352	8	US-08-979-847-118	Sequence 118, App
21	33	82.5	378	8	US-08-979-847-122	Sequence 122, App
22	33	82.5	395	15	US-10-295-403-114	Sequence 114, App
23	33	82.5	398	8	US-08-979-847-121	Sequence 121, App
24	33	82.5	420	9	US-09-864-761-37994	Sequence 37994, A
25	33	82.5	488	9	US-09-840-787-17	Sequence 17, Appl
26	33	82.5	577	14	US-10-072-841-29	Sequence 29, Appl
27	33	82.5	798	10	US-09-978-249-8	Sequence 8, Appl
28	33	82.5	813	15	US-10-105-989-2	Sequence 2, Appl
29	32	80.0	54	10	US-09-936-194-4	Sequence 4, Appl
30	32	80.0	77	10	US-09-764-864-1029	Sequence 1029, Ap
31	32	80.0	86	9	US-09-864-761-43075	Sequence 43075, A
32	32	80.0	119	10	US-09-764-864-1485	Sequence 1485, Ap
33	32	80.0	119	10	US-09-764-847-768	Sequence 768, App
34	32	80.0	119	15	US-10-092-154-768	Sequence 768, App
35	32	80.0	120	10	US-09-764-864-970	Sequence 970, App
36	32	80.0	132	10	US-09-764-864-1405	Sequence 1405, Ap
37	32	80.0	148	15	US-10-106-698-5950	Sequence 5950, Ap
38	32	80.0	149	15	US-10-156-932-28	Sequence 28, Appl
39	32	80.0	156	10	US-09-764-864-1319	Sequence 1319, Ap
40	32	80.0	156	15	US-10-106-698-6111	Sequence 6111, Ap
41	32	80.0	168	10	US-09-764-864-919	Sequence 919, App
42	32	80.0	175	10	US-09-764-864-875	Sequence 875, App
43	32	80.0	175	10	US-09-764-864-1326	Sequence 1326, Ap
44	32	80.0	184	10	US-09-764-864-938	Sequence 938, Ap
45	32	80.0	187	10	US-09-764-864-1361	Sequence 1361, Ap

ALIGNMENTS

RESULT 1

US-10-057-890A-6
; Sequence 6, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same
; TITLE OF INVENTION: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-6

Query Match 100.0% Score 40; DB 15; Length 6;
Best Local Similarity 100.0%; Pred No. 5,3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YKGLC 6
| | | | |
Db 1 YKGLC 6
RESULT 2
US-10-057-890A-10
; Sequence 10, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy

```
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, and
; FILE REFERENCE: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 10
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-10

Query Match      100.0%; Score 40; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      31 YKCGLC 36

RESULT 3
US-10-057-890A-31
; Sequence 31, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, and
; FILE REFERENCE: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 31
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-31

Query Match      100.0%; Score 40; DB 15; Length 157;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      50 YKCGLC 55

RESULT 4
US-09-949-029-18
; Sequence 18, Application US/09949029
; Publication No. US20030134278A1
; GENERAL INFORMATION:
; APPLICANT: Karpen, G.H.
; APPLICANT: Dobie, K.W.
; APPLICANT: Kennedy, C.D.
; APPLICANT: Velasco, V.M.
; APPLICANT: McGrath, T.L.
; APPLICANT: Weko, J.
; APPLICANT: Patterson, R.W.
; TITLE OF INVENTION: Identification of chromosome inheritance modifiers in Drosophila
; FILE REFERENCE: 1211.015US1
```

```
; CURRENT APPLICATION NUMBER: US/09/949,029
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/231,178
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3067
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-949-029-18

Query Match      90.0%; Score 36; DB 12; Length 3067;
Best Local Similarity 83.3%; Pred. No. 8.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      2813 YKCGLC 2818

RESULT 5
US-09-867-550-760
; Sequence 760, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: NO. US20020082206A1el Polynucleotides from Atherogenic Cells
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 760
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-760

Query Match      87.5%; Score 35; DB 9; Length 250;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      82 YKCGAC 87

RESULT 6
US-09-800-729-196
; Sequence 196, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 196
; LENGTH: 313
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-196

Query Match      87.5%; Score 35; DB 9; Length 313;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      208 YKCGAC 213

RESULT 7
US-09-764-864-967
; Sequence 967, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Bae, Kwang-Hee
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Kwon, Young Do
; APPLICANT: Ryu, Eun-Hyun
; APPLICANT: Hwang, Moon-Sun
; TITLE OF INVENTION: ZINC FINGER DOMAIN LIBRARIES
; FILE REFERENCE: 12279-005001
; CURRENT APPLICATION NUMBER: US/10/223,765
; CURRENT FILING DATE: 2001-08-19
; PRIOR APPLICATION NUMBER: 60/374,355
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 60/313,402
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-765-39

Query Match      85.0%; Score 34; DB 12; Length 23;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      1 YKCGQC 6

RESULT 10
US-09-764-864-927
; Sequence 927, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 927
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (200)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-864-927

Query Match      85.0%; Score 34; DB 10; Length 308;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      1 YKCGQC 6
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-196

Query Match      87.5%; Score 35; DB 9; Length 316;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      192 YKCGLC 197

RESULT 8
US-09-785-632A-39
; Sequence 39, Application US/09785632A
; Patent No. US20020061512A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo
; APPLICANT: Kwon, Young Do
; APPLICANT: Kim, Hyun-Won
; APPLICANT: Ryu, Eun-Hyun
; APPLICANT: Hwang, Moon-Sun
; TITLE OF INVENTION: ZINC FINGER DOMAINS AND METHODS OF
IDENTIFYING SAME
; FILE REFERENCE: 12279-002001
; CURRENT APPLICATION NUMBER: US/09/785,632A
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-785-632A-39

Query Match      85.0%; Score 34; DB 9; Length 23;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      1 YKCGQC 6
```

```

Db      242 YKCNLC 247
      ||| ||
RESULT 11
US-09-764-864-1366
; Sequence 1366, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1366
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1366

Query Match      85.0%; Score 34; DB 10; Length 309;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      243 YKCNLC 248
      ||| ||
RESULT 12
US-09-949-029-96
; Sequence 96, Application US/09949029
; Publication No. US20030134278A1
; GENERAL INFORMATION:
; APPLICANT: Karpen, G.H.
; APPLICANT: Dobie, K.W.
; APPLICANT: Kennedy, C.D.
; APPLICANT: Velasco, V.M.
; APPLICANT: McGrath, T.L.
; APPLICANT: Weko, J.
; APPLICANT: Patterson, R.W.
; TITLE OF INVENTION: Identification of chromosome inheritance modifiers in Drosophila
; TITLE OF INVENTION: melanogaster
; FILE REFERENCE: 1211.015US1
; CURRENT APPLICATION NUMBER: US/09/949,029
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/231,178
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-949-029-96

Query Match      85.0%; Score 34; DB 12; Length 366;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      158 YKCGSC 163
      |||| |
RESULT 13
US-09-372-348-11
; Sequence 11, Application US/09372348A
; Publication No. US20030028005A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Hedrick, Joseph A.
; APPLICANT: Wang, Luquan
; APPLICANT: Zlotnik, Albert
; APPLICANT: Murgolo, Nicholas J.
; APPLICANT: Greene, Jonathan R.
; APPLICANT: Johnston, James A.
; APPLICANT: Bazan, J. Fernando
; TITLE OF INVENTION: Mammalian Proteins; Related Reagents and Methods
; FILE REFERENCE: DX084K(2d)
; CURRENT APPLICATION NUMBER: US/09/372,348A
; CURRENT FILING DATE: 1999-08-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 530
; TYPE: PRT
; ORGANISM: primate
US-09-372-348-11

Query Match      85.0%; Score 34; DB 11; Length 530;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      252 YKCGLC 257
      |||| |
RESULT 14
US-09-764-864-957
; Sequence 957, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 957
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-957

Query Match      85.0%; Score 34; DB 10; Length 600;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db      380 YKCALC 385
      ||| ||
RESULT 15
US-09-801-368-34
; Sequence 34, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff

```

; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 1323
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-34

Query Match 85.0%; Score 34; DB 10; Length 1323;
Best Local Similarity 83.3%; Pred. No. 8.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YKCGLC 6
|
Db 132 YPCGIC 137

Search completed: October 1, 2003, 10:50:53
Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:33:32 ; Search time 35.8333 Seconds
(without alignments)
43.209 Million cell updates/sec

Title: US-10-057-890A-6
Perfect score: 40
Sequence: 1 YKGLC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_23:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	127	4 Q9BTS6	Q9BTS6 homo sapien
2	40	100.0	366	11 Q8C1M2	Q8C1M2 mus musculus
3	40	100.0	433	5 Q95X98	Q95X98 caenorhabdi
4	40	100.0	601	11 Q8BJ07	Q8BJ07 mus musculus
5	38	95.0	190	5 Q8SW43	Q8SW43 encephalito
6	38	95.0	357	4 Q96JW6	Q96JW6 homo sapien
7	38	95.0	406	4 Q96MW3	Q96MW3 homo sapien
8	38	95.0	503	11 Q921H7	Q921H7 mus musculus
9	38	95.0	591	11 Q8R0V0	Q8R0V0 mus musculus
10	38	95.0	618	4 Q96MX3	Q96MX3 homo sapien
11	38	95.0	652	4 Q9NS42	Q9NS42 homo sapien
12	38	95.0	714	4 Q8TE95	Q8TE95 homo sapien
13	38	95.0	769	4 Q81ZD3	Q81ZD3 homo sapien
14	38	95.0	837	5 Q9NAS7	Q9NAS7 anopheles g
15	37	92.5	57	4 Q15920	Q15920 homo sapien
16	37	92.5	149	16 Q97RB5	Q97RB5 streptococc

17	37	92.5	149	16	Q8DQ78	Q8DQ78 streptococc
18	37	92.5	193	10	Q8L8P4	Q8L8P4 arabidopsis
19	37	92.5	193	10	Q9SSW0	Q9SSW0 arabidopsis
20	37	92.5	228	10	Q8H5V7	Q8H5V7 oryza sativ
21	37	92.5	337	11	Q8BG98	Q8BG98 mus musculu
22	37	92.5	451	11	Q8B1K0	Q8B1K0 mus musculu
23	37	92.5	522	4	Q8N9F8	Q8N9F8 homo sapien
24	37	92.5	538	5	Q9VGG0	Q9VGG0 drosophila
25	37	92.5	626	4	Q9BX82	Q9BX82 homo sapien
26	37	92.5	751	4	Q9NS43	Q9NS43 homo sapien
27	37	92.5	1004	5	Q8SZM4	Q8SZM4 drosophila
28	36	90.0	155	10	Q8H7J9	Q8H7J9 oryza sativ
29	36	90.0	348	10	Q8GXX6	Q8GXX6 arabidopsis
30	36	90.0	348	13	Q93384	Q93384 oncorhynch
31	36	90.0	360	11	Q8VHM2	Q8VHM2 mus musculu
32	36	90.0	360	11	Q70600	Q70600 rattus norv
33	36	90.0	361	4	Q8WV14	Q8WV14 homo sapien
34	36	90.0	361	4	Q8WVG1	Q8WVG1 homo sapien
35	36	90.0	362	6	Q9WZU4	Q9WZU4 sus scrofa
36	36	90.0	362	11	Q9JHD4	Q9JHD4 mus musculu
37	36	90.0	362	11	Q8CBB9	Q8CBB9 mus musculu
38	36	90.0	388	5	Q8SUH6	Q8SUH6 encephalito
39	36	90.0	581	11	Q8C7A0	Q8C7A0 mus musculu
40	36	90.0	1638	5	O45853	O45853 caenorhabdi
41	36	90.0	3313	5	Q8W3W6	Q8W3W6 drosophila
42	35	87.5	102	17	Q8TQ05	Q8TQ05 methanosarc
43	35	87.5	130	6	Q95286	Q95286 sus scrofa
44	35	87.5	183	9	Q8LrG6	Q8LrG6 streptococc
45	35	87.5	275	11	Q35480	Q35480 cricetus

ALIGNMENTS

RESULT 1

Q9BTS6 PRELIMINARY; PRT; 127 AA.
ID Q9BTS6;
AC Q9BTS6;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DI 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003386; AAH03386.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 127 AA; 14474 MW; 24C29EA7A20FABD1 CRC64;

Query Match 100.0%; Score 40; DB 4; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
DB 65 YKGLC 70

RESULT 2

Q8C1M2 PRELIMINARY; PRT; 366 AA.
ID Q8C1M2;
AC Q8C1M2;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Similar to zinc finger protein 29.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH 11;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC023282; AAH23282.1; -.
 SQ SEQUENCE 366 AA; 41293 MW; 3C4B50C8A73700B6 CRC64;

Query Match 100.0%; Score 40; DB 11; Length 366;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 117 YKCGLC 122

RESULT 3
 Q95X98 PRELIMINARY; PRT; 433 AA.
 AC Q95X98;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 49.8 kDa protein.
 GN Y37F4.6
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peleoderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.*;
 Science 282:2012-2018(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.*;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC098856; AAL27253.1; -.
 DR WormPep; Y37F4.6; CE29854.
 DR InterPro; IPR007087; znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 1.
 DR SMART; SM00355; znf_C2H2; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 433 AA; 49833 MW; 9F373420CBDE834D CRC64;

Query Match 100.0%; Score 40; DB 5; Length 433;
 Best Local Similarity 100.0%; Pred. No. 9.6;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 93 YKCGLC 98

RESULT 4
 Q8BJ07 PRELIMINARY; PRT; 601 AA.
 AC Q8BJ07;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Zinc finger protein 31.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.*;
 Nature 420:563-573(2002).
 RL Nature 420:563-573(2002).
 DR EMBL; AK047820; BAC33164.1; -.
 SQ SEQUENCE 601 AA; 67542 MW; 90B65BF55B117C08 CRC64;

Query Match 100.0%; Score 40; DB 11; Length 601;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 352 YKCGLC 357

RESULT 5
 Q8SW43 PRELIMINARY; PRT; 190 AA.
 AC Q8SW43;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein ECU03_0790.
 GN ECU03_0790.
 OS Encephalitozoon cuniculi.
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
 OX NCBI_TaxID=6035;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RA Genoscope;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GB-M1;
 RX MEDLINE=21576510; PubMed=11719806;
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarot F.,
 RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
 RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
 RA Weissenbach J., Vivares C.P.;
 RT "Genome sequence and gene compaction of the eukaryote parasite
 Encephalitozoon cuniculi.*;
 Nature 414:450-453(2001).
 RL Nature 414:450-453(2001).
 DR EMBL; AL590443; CAD26223.1; -.
 DR InterPro; IPR007087; znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 6.
 DR SMART; SM00355; znf_C2H2; 5.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 190 AA; 22245 MW; C2006BC95F3AA5B1 CRC64;

Query Match 95.0%; Score 38; DB 5; Length 190;
 Best Local Similarity 83.3%; Pred. No. 12;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 85 YKCGLC 90

RESULT 6

Q96JW6 PRELIMINARY; PRT: 357 AA.
ID Q96JW6
AC Q96JW6
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ14928.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Niimiyama K., Iwayanagi T.;
RT *NEO human cDNA sequencing project.*;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AK027834; BAB55400.1; -!
DR InterPro: IPR007087; Znf_C2H2.
DR InterPro: IPR007086; Znf_C2H2_sub.
DR Pfam: PF00096; zf_C2H2; 12.
DR PRINTS: PR00048; ZINCFINGER.
DR PRODOM: PD000003; Znf_C2H2; 1.
DR SMART: SM00355; Znf_C2H2; 13.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 13.
KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
KW Zinc-finger.
SQ SEQUENCE 357 AA; 40426 MW; 541DBE6D009B8C04 CRC64;

Query Match 95.0%; Score 38; DB 4; Length 357;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
| | | | |

Db 334 YKCGIC 339

RESULT 7

Q96MV3 PRELIMINARY; PRT: 406 AA.
ID Q96MV3
AC Q96MV3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ31827.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekino M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT *NEO human cDNA sequencing project.*;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK056389; BAB71174.1; -!
DR InterPro: IPR001909; KRAB.

DR InterPro: IPR007087; Znf_C2H2.
DR Pfam: PF00096; zf_C2H2; 10.
DR SMART: SM00349; KRAB; 1.
DR SMART: SM00355; Znf_C2H2; 10.
DR PROSITE: PS0805; KRAB; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 10.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 12.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 406 AA; 46686 MW; 1BA9BC5DF90886 CRC64;

Query Match 95.0%; Score 38; DB 4; Length 406;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
| | | | |

Db 383 YKCGIC 388

RESULT 8

Q92LH7 PRELIMINARY; PRT: 563 AA.
ID Q92LH7
AC Q92LH7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Unknown (Protein for IMAGE:3708275) (Fragment).
GN BC026401.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC012403; AAH12403.1; -!
DR MGI: 2384725; BC026401.
DR InterPro: IPR007087; Znf_C2H2.
DR Pfam: PF00096; zf_C2H2; 12.
DR SMART: SM00355; Znf_C2H2; 12.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 12.
KW Metal-binding; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 563 AA; 61463 MW; D446AD9E8ED7C45F CRC64;

Query Match 95.0%; Score 38; DB 11; Length 563;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
| | | | |

Db 163 YKCGIC 168

RESULT 9

Q8ROV0 PRELIMINARY; PRT: 591 AA.
ID Q8ROV0
AC Q8ROV0
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (Hypothetical zinc finger).
GN BC026401.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=22354683; PubMed=12466851;
 RA THE FANTOM Consortium,
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 DR EMBL; AC026401; AAH26401.1; -;
 DR EMBL; AK049208; BAC33610.1; -;
 DR MGD; MGI:2384725; BC026401.
 DR InterPro; IPR007087; znf_C2H2.
 DR Pfam; PF00096; znf_C2H2; 12.
 DR SMART; SM00355; znf_C2H2; 12.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 12.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 591 AA; 64599 MW; EE070D2692FA4A73 CRC64;

Query Match 95.0%; Score 38; DB 11; Length 591;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 191 YKCGIC 196

RESULT 10

Q96MX3
 ID Q96MX3 PRELIMINARY; PRT; 618 AA.
 AC Q96MX3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ31751.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
 RA Hotata T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
 RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RT "NEO human cDNA sequencing project";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK056313; BAB11146.1; -;
 DR InterPro; IPR007087; znf_C2H2.
 DR Pfam; PF00096; znf_C2H2; 12.
 DR SMART; SM00355; znf_C2H2; 12.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 12.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 618 AA; 67832 MW; 28C0F13E54773353 CRC64;

Query Match 95.0%; Score 38; DB 4; Length 618;
 Best Local Similarity 83.3%; Pred. No. 30;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 220 YKCGIC 225

RESULT 11

-

Q9NS42
 ID Q9NS42 PRELIMINARY; PRT; 652 AA.
 AC Q9NS42;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical zinc finger-like protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kodoylanni V., Ge Y., Severin J., Krummel G.K., Grable L.,
 RA Kviststad E., Gordon L., Shannon M., Brower A., Olsen A.S., Smith L.M.;
 RT "Sequence analysis of a 1MD region in 19q13.2 containing a zinc finger
 gene cluster";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL; AC074331; AAF88106.1; -;
 DR HSSP; P08045; 12NF.
 DR InterPro; IPR007087; znf_C2H2.
 DR InterPro; IPR007086; znf_C2H2_sub.
 DR InterPro; IPR006025; znf_MTpeptdse.
 DR Pfam; PF00096; znf_C2H2; 8.
 DR PRINTS; PD00048; ZINC_FINGER.
 DR ProDom; PD000003; znf_C2H2; 7.
 DR SMART; SM00355; znf_C2H2; 7.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 8.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
 SQ SEQUENCE 652 AA; 72644 MW; 3C9F250594170178 CRC64;

Query Match 95.0%; Score 38; DB 4; Length 652;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 602 YKCGMC 607

RESULT 12

Q8TE95
 ID Q8TE95 PRELIMINARY; PRT; 714 AA.
 AC Q8TE95;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ23765.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Mishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEO human cDNA sequencing project";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK074345; BAB85057.1; -;
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR007087; znf_C2H2.
 DR Pfam; PF00096; znf_C2H2; 21.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; znf_C2H2; 21.
 DR PROSITE; PS50805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 21.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 23.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 714 AA; 81668 MW; EFD0E080DE78FC06 CRC64;

Query Match 95.0%; Score 38; DB 4; Length 714;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 691 YKCGIC 696
 |||||

RESULT 13
 Q81ZD3 PRELIMINARY; PRT; 769 AA.
 AC Q81ZD3;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE zinc finger protein 41-like protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA Guo J.H., Yu L.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY147015; AAN61169.1; -.
 SQ SEQUENCE 769 AA; 88000 MW; CC670B86FFCABBFD CRC64;

Query Match 95.0%; Score 38; DB 4; Length 769;
 Best Local Similarity 83.3%; Pred. No. 36;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 746 YKCGIC 751
 |||||

RESULT 14
 Q9NAS7 PRELIMINARY; PRT; 837 AA.
 AC Q9NAS7;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Integrin beta subunit precursor.
 GN BINT.
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
 OX NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3;
 RX MEDLINE=21331701; PubMed=11437913;
 RA Machairaki V., Lycett G., Blass C., Louis C.;
 RT "Beta-Integrin of Anopheles gambiae: mRNA cloning and analysis of
 RT structure and expression."
 RL Insect Mol. Biol. 10:217-223(2001).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- PTM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE
 CC BONDS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 DR EMBL; AJ292755; CAC00630.1; -.
 DR HSSP; P05106; 1JY2
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR002369; Integrin_B.
 DR InterPro; IPR001169; Integrin_beta_C.
 DR InterPro; IPR003659; Plexin-like.
 DR Pfam; PF00362; Integrin_B.1.
 DR PRINTS; PRO1186; INTEGRINB.

DR ProDom: PD001811; Integrin_B; 1.
 DR SMART; SM00187; INB; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00243; INTEGRIN_BETA; 3.
 KW Cell adhesion; Glycoprotein; Integrin; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 26
 FT CHAIN 27 837 BETA INTEGRIN SUBUNIT.
 SQ SEQUENCE 837 AA; 92791 MW; 30F1F339D6924D78 CRC64;

Query Match 95.0%; Score 38; DB 5; Length 837;
 Best Local Similarity 83.3%; Pred. No. 38;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 518 YKCGIC 523
 |||||

RESULT 15
 Q15920 PRELIMINARY; PRT; 57 AA.
 AC Q15920;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE DNA-binding protein (fragment).
 GN ZNF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=92052132; PubMed=1946370;
 RA Bray P., Lichter P., Thiesen H.J., Ward D.C., David I.B.;
 RT "Characterization and mapping of human genes encoding zinc finger
 RT proteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9563-9567(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=92372070; PubMed=1505991;
 RA Lichter P., Bray P., Ried T., David I.B., Ward D.C.;
 RT "Clustering of C2-H2 zinc finger motifs sequences within telomeric and
 RT fragile site regions of human chromosomes.";
 RL Genomics 13:999-1007(1992).
 DR EMBL; M88358; AAA61316.1; -.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 1.
 DR SMART; SM0355; Znf_C2H2; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
 KW DNA-binding; Metal-binding; Zinc; Zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 57 AA; 6375 MW; 9A482948BAB9429A CRC64;

Query Match 92.5%; Score 37; DB 4; Length 57;
 Best Local Similarity 83.3%; Pred. No. 6.9;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 15 YKCGVC 20
 |||||

Search completed: October 1, 2003, 10:46:29
 Job time : 38.8333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:06:46 ; Search time 7 Seconds
(without alignments)
40.309 Million cell updates/sec

Title: US-10-057-890A-6
Perfect score: 40
Sequence: 1 YKGLC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	427	1	MURE_CAMJE	O69290 campylobact
2	40	100.0	614	1	2F29_MOUSE	Q07230 mus musculu
3	40	100.0	1350	1	XFIN_XENLA	Q06045 xenopus lae
4	36	90.0	581	1	2319_MOUSE	Q9er18 mus musculu
5	36	90.0	582	1	2319_HUMAN	Q9p2f9 homo sapien
6	35	87.5	682	1	2N45_HUMAN	Q02386 homo sapien
7	35	87.5	722	1	2219_HUMAN	Q9p2v4 homo sapien
8	35	87.5	1115	1	YAA4_SCHPO	Q09798 schizosacch
9	34	85.0	235	1	OYOL_HUMAN	Q14753 homo sapien
10	34	85.0	267	1	OYOL_MOUSE	Q9wtj2 mus musculu
11	34	85.0	352	1	2185_MOUSE	Q62394 mus musculu
12	34	85.0	457	1	2185_HUMAN	Q15231 homo sapien
13	34	85.0	472	1	L126_CABEL	Q27355 caenorhabdi
14	34	85.0	495	1	MLP2_DROME	Q24400 drosophila
15	34	85.0	506	1	Y619_METJA	Q58036 methanococc
16	34	85.0	514	1	2BT2_HUMAN	Q8n680 homo sapien
17	34	85.0	605	1	TC17_HUMAN	Q60765 homo sapien
18	34	85.0	798	1	17B5_MOUSE	Q07309 mus musculu
19	34	85.0	1028	1	OYO_DROME	P51521 drosophila
20	34	85.0	1323	1	ADRI_YEAST	P07248 saccharomyc
21	33	82.5	201	1	2239_MOUSE	Q24399 mus musculu
22	33	82.5	223	1	GLI4_HUMAN	P10075 homo sapien
23	33	82.5	273	1	2N80_GORGO	P31503 gorilla gor
24	33	82.5	293	1	2N80_MACMU	P31505 macaca mula
25	33	82.5	407	1	OZF_MOUSE	Q62513 mus musculu
26	33	82.5	407	1	OZF_RAT	Q62981 rattus norv
27	33	82.5	458	1	2239_HUMAN	Q16600 homo sapien
28	33	82.5	577	1	17B6_CAVPO	P18563 cavia porce
29	33	82.5	634	1	2297_HUMAN	Q15209 homo sapien
30	33	82.5	645	1	2235_MOUSE	Q61116 mus musculu
31	33	82.5	655	1	17B5_PAPCY	Q07441 papio cynoc
32	33	82.5	697	1	Y441_HUMAN	O43167 homo sapien
33	33	82.5	698	1	2234_HUMAN	Q14588 homo sapien

34	33	82.5	744	1	YJ62_HUMAN	Q8tf39 homo sapien
35	33	82.5	744	1	2366_HUMAN	Q88895 homo sapien
36	33	82.5	799	1	17B5_HUMAN	P18084 homo sapien
37	33	82.5	803	1	2226_HUMAN	Q9nyt6 homo sapien
38	33	82.5	918	1	PEP3_YEAST	P27801 saccharomyc
39	33	82.5	1191	1	ZN91_HUMAN	Q05481 homo sapien
40	32	80.0	36	1	DESR_DESGI	P00273 desulfovibr
41	32	80.0	294	1	FDNH_ECOLI	P24184 escherichia
42	32	80.0	300	1	FDH_ECOLI	P32175 escherichia
43	32	80.0	312	1	FDXH_HAEIN	P44450 haemophilus
44	32	80.0	405	1	VIB2_NPVOV	P32511 orgyia pseu
45	32	80.0	416	1	CREA_EMENI	Q01981 emericeella

ALIGNMENTS

RESULT 1
MURE_CAMJE
ID MURE_CAMJE STANDARD; PRI; 427 AA.
AC O69290: O9PM35;
DT 15-JUL-1999 (Rel. 38, Created)
DI 16-OCT-2001 (Rel. 40, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE UDP-N-acetylmuramoyl-D-glutamate--2,6-diaminopimelate ligase
(EC 6.3.2.13) (UDP-N-acetylmuramyl-tripeptide synthetase) (Meso-
diaminopimelate-adding enzyme) (UDP-MurNAC-tripeptide synthetase).
GN MURE OR CJ1641.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
Whitehead S., Barrrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
reveals hypervariable sequences.";
RN Nature 403:665-668(2000).
[2]
RC SEQUENCE OF 1-326 FROM N.A.
RA STRAIN=NCTC 11168;
RL Griffiths P.L., Connerton I.F.;
CC Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CELL WALL FORMATION. DIAMINOPIMELIC ACID ADDING ENZYME
(BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanyl-D-
glutamate + meso-2,6-diaminoheptanedioate - ADP + phosphate + UDP-
N-acetylmuramoyl-L-alanyl-D-gamma-glutamyl-meso-2,6-diamino-
heptanedioate
CC -1- PATHWAY: Peptidoglycan biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the murCDEF family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: AL139079; CAB73629.1; -;
DR EMBL: Y16862; CAA76495.1; ALT_INIT.
DR PIR: G81260; G81260.
DR HAMAP: MF_00208; -; 1.
DR InterPro: IPR000713; Mur_ligase.
DR InterPro: IPR004101; Mur_ligase_C.
DR

DR InterPro: IPR005761; Mure.
DR Pfam: PF01225; Mur_ligase; 1.
DR Pfam: PF02875; Mur_ligase_C; 1.
DR TIGRfams: TIGR01089; mure; 1.
KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
KW ATP-binding; Complete proteome.
FT NP_BIND 66 72 ATP (POTENTIAL).
FT CONFLICT 1 3 MKL -> EIK (IN REF. 2).
FT CONFLICT 39 39 K -> G (IN REF. 2).
FT CONFLICT 54 54 K -> E (IN REF. 2).
FT CONFLICT 121 121 Q -> L (IN REF. 2).
FT CONFLICT 125 125 Q -> P (IN REF. 2).
FT CONFLICT 134 134 E -> G (IN REF. 2).
FT CONFLICT 142 142 Q -> P (IN REF. 2).
FT CONFLICT 145 145 I -> S (IN REF. 2).
FT CONFLICT 166 166 H -> R (IN REF. 2).
FT CONFLICT 170 170 E -> G (IN REF. 2).
SQ SEQUENCE 427 AA; 47721 MW; FADBFIE67471B824 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
DB 86 YKGLC 91

RESULT 2
ZF29_MOUSE
ID ZF29_MOUSE STANDARD; PRT; 614 AA.
AC Q07230;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 29 (Zfp-29).
GN ZFP29 OR ZFP-29.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RA MEDLINE=92039080; PubMed=1937051;
RA Denny P.; Ashworth A.;
RT "A zinc finger protein"-encoding gene expressed in the post-meiotic
RT phase of spermatogenesis."
RL Gene 106.221-227(1991).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION DURING
CC THE POST-MEIOTIC STAGES OF SPERMATOGENESIS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: IN THE ADULT, PREDOMINANTLY FOUND IN
CC SPERMATIDS. ALSO PRESENT IN THE EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 1 SCAN box domain.
CC -!- SIMILARITY: Contains 14 C2H2-type zinc fingers.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X55126; CAA38920.1; -.
DR PIR: JH0500; JH0500.
DR HSP: P08047; 1SP2.
DR MGD: MGI:99176; Zfp29.
DR InterPro: IPR003309; Treg_SCAN.
DR InterPro: IPR007087; Znf_C2H2.

DR InterPro: IPR007086; Znf_C2H2_sub.
DR Pfam: PF02023; SCAN; 1.
DR Pfam: PF00096; Zf_C2H2; 14.
DR PRINTS: PR00048; ZINCFINGER.
DR PRODOM: PD000003; Znf_C2H2; 13.
DR SMART: SM00431; LER; 1.
DR SMART: SM00355; Znf_C2H2; 14.
DR PROSITE: PS00804; SCAN_BOX; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 14.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 14.
KW Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat;
KW Developmental protein; Transcription regulation; Spermatogenesis.
FT DOMAIN 69 127 SCAN_BOX.
FT ZN_FING 222 244 C2H2-TYPE 1.
FT ZN_FING 250 272 C2H2-TYPE 2.
FT ZN_FING 278 300 C2H2-TYPE 3.
FT ZN_FING 306 328 C2H2-TYPE 4.
FT ZN_FING 334 356 C2H2-TYPE 5.
FT ZN_FING 362 384 C2H2-TYPE 6.
FT ZN_FING 390 412 C2H2-TYPE 7.
FT ZN_FING 418 440 C2H2-TYPE 8.
FT ZN_FING 446 468 C2H2-TYPE 9.
FT ZN_FING 474 496 C2H2-TYPE 10.
FT ZN_FING 502 524 C2H2-TYPE 11.
FT ZN_FING 530 552 C2H2-TYPE 12.
FT ZN_FING 558 580 C2H2-TYPE 13.
FT ZN_FING 586 608 C2H2-TYPE 14.
SQ SEQUENCE 614 AA; 58714 MW; 738156A0A9413DF7 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 614;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
DB 446 YKGLC 451

RESULT 3
XF1N_XENLA
ID XF1N_XENLA STANDARD; PRT; 1350 AA.
AC P08045;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein Xfin.
GN XF1N.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Embryo;
RA MEDLINE=88082679; PubMed=2826129;
RA Ruiz i Altaba A.; Perry-O'Keefe H.; Melton D.A.;
RT "Xfin: an embryonic gene encoding a multifingered protein in
RT Xenopus."
RL EMBO J. 6:3065-3070(1987).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94021366; PubMed=7692399;
RA Andreazoli M.; de Lucchini S.; Costa M.; Barsacchi G.;
RT "RNA binding properties and evolutionary conservation of the Xenopus
RT multifinger protein Xfin."
RL Nucleic Acids Res. 21:4218-4225(1993).
RN [3]
RP STRUCTURE BY NMR OF FINGER 31.
RX MEDLINE=89346749; PubMed=2503871;
RA Lee M.S.; Gippert G.P.; Soman K.V.; Case D.A.; Wright P.E.;
RT "Three-dimensional solution structure of a single zinc finger DNA-
RT binding domain.";

Science 245:635-637(1989).

[4]

RP MEDLINE-89378224; PubMed-2506074;

RA Lee M.S., Cavanagh J., Wright P.E.;

RT "Complete assignment of the 1H NMR spectrum of a synthetic zinc

RT finger from xfin. Sequential resonance assignments and secondary

RT structure";

RL FEBS Lett. 254:159-164(1989).

CC -!- FUNCTION: BINDS RNA. COULD FUNCTION IN POST-TRANSLATIONAL

CC REGULATION PROCESSES.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- TISSUE SPECIFICITY: OOCYTES AND IN SPECIALIZED CELL TYPES SUCH

CC AS NEURAL RETINA CONES.

CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT OOGENESIS AND

CC EMBRYOGENESIS.

CC -!- DOMAIN: CONTAINS 37 FINGER MOTIFS IN 6 DOMAINS.

CC -!- PTM: PHOSPHORYLATED.

CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-

CC FINGER PROTEINS.

CC -!- SIMILARITY: Contains 1 KRAB domain.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; X06021; CAA29425.1; -

DR PDB; 1ZNF; 15-OCT-91.

DR InterPro: IPR001909; KRAB.

DR InterPro: IPR007087; Znf_C2H2.

DR Pfam: PF01352; KRAB; 1.

DR Pfam: PF00096; zf-C2H2; 36.

DR ProDom; PD000003; Znf_C2H2; 20.

DR SMART; SM00349; KRAB; 1.

DR SMART; SM00355; Znf_C2H2; 35.

DR PROSITE; PS0805; KRAB; 1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 35.

DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 37.

DR Zinc-finger; Metal-binding; RNA-binding; Repeat; 3D-structure;

KW Phosphorylation.

KW

FT DOMAIN 1 58 KRAB.

FT ZN_FING 108 130 C2H2-TYPE.

FT ZN_FING 136 158 C2H2-TYPE.

FT ZN_FING 164 186 C2H2-TYPE.

FT ZN_FING 192 214 C2H2-TYPE.

FT ZN_FING 220 242 C2H2-TYPE.

FT ZN_FING 248 270 C2H2-TYPE.

FT ZN_FING 276 298 C2H2-TYPE.

FT ZN_FING 326 348 C2H2-TYPE.

FT ZN_FING 354 376 C2H2-TYPE.

FT ZN_FING 382 404 C2H2-TYPE.

FT ZN_FING 410 432 C2H2-TYPE.

FT ZN_FING 438 460 C2H2-TYPE.

FT ZN_FING 466 488 C2H2-TYPE.

FT ZN_FING 503 525 C2H2-TYPE.

FT ZN_FING 531 553 C2H2-TYPE.

FT ZN_FING 559 581 C2H2-TYPE.

FT ZN_FING 587 609 C2H2-TYPE.

FT ZN_FING 615 637 C2H2-TYPE.

FT ZN_FING 643 665 C2H2-TYPE.

FT ZN_FING 671 693 C2H2-TYPE.

FT ZN_FING 699 721 C2H2-TYPE.

FT ZN_FING 750 772 C2H2-TYPE.

FT ZN_FING 778 800 C2H2-TYPE.

FT ZN_FING 806 828 C2H2-TYPE.

FT ZN_FING 834 856 C2H2-TYPE.

FT ZN_FING 862 884 C2H2-TYPE.

FT ZN_FING 890 912 C2H2-TYPE.

FT ZN_FING 918 940 C2H2-TYPE.

FT ZN_FING 988 1010 C2H2-TYPE.

FT ZN_FING 1016 1038 C2H2-TYPE.

FT ZN_FING 1044 1066 C2H2-TYPE.

FT ZN_FING 1136 1158 C2H2-TYPE.

FT ZN_FING 1164 1186 C2H2-TYPE.

FT ZN_FING 1192 1214 C2H2-TYPE.

FT ZN_FING 1220 1242 C2H2-TYPE.

FT ZN_FING 1248 1270 C2H2-TYPE.

FT ZN_FING 1276 1298 C2H2-TYPE.

FT STRAND 1045 1045

FT STRAND 1052 1052

FT HELIX 1056 1062

FT HELIX 1063 1065

FT TURN 1066 1066

SQ SEQUENCE 1350 AA; 155804 MW; 27F10AB0851E0AD8 CRC64;

Query Match 100.0%; Score 40; DB 1; Length 1350;

Best Local Similarity 100.0%; Pred. No. 8.8;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLGC 6

Db 1044 YKGLGC 1049

RESULT 4

ID Z319_MOUSE STANDARD; PRT; 581 AA.

AC Q9ER88; 2003 (Rel. 41, Created)

DI 28-FEB-2003 (Rel. 41, Last sequence update)

DI 15-SEP-2003 (Rel. 42, Last annotation update)

DE Zinc finger protein 319.

GN ZNF319 OR ZFP319.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-21100867; PubMed-11161788;

RA Laub F., Aldabe R., Ou J., Ramirez F.;

RT "Overexpression of novel zinc-finger protein induces apoptosis in

RT NIH3T3 fibroblasts.";

RL Genomics 70:375-380(2000).

CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-

CC FINGER PROTEINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; AF288403; AAG28743.1; -

CC HSPB; P08045; 12NF.

CC MGD; MGI:1890618; zfp319.

CC GO; GO:0005634; C.nucleus; IDA.

CC InterPro: IPR007087; Znf_C2H2.

CC Pfam; PF00096; zf-C2H2; 12.

CC ProDom; PD000003; Znf_C2H2; 4.

CC SMART; SM00355; Znf_C2H2; 12.

CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.

CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 15.

KW Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;

FT ZN_FING 75 99 C2H2-TYPE.

FT ZN_FING 103 125 C2H2-TYPE (DEGENERATE).

FT ZN_FING 131 153 C2H2-TYPE.

FT ZN_FING 201 223 C2H2-TYPE.
 FT ZN_FING 229 251 C2H2-TYPE.
 FT ZN_FING 257 279 C2H2-TYPE.
 FT ZN_FING 286 308 C2H2-TYPE (DEGENERATE).
 FT ZN_FING 314 336 C2H2-TYPE.
 FT ZN_FING 342 364 C2H2-TYPE.
 FT ZN_FING 370 392 C2H2-TYPE (DEGENERATE).
 FT ZN_FING 398 420 C2H2-TYPE.
 FT ZN_FING 427 449 C2H2-TYPE.
 FT ZN_FING 457 479 C2H2-TYPE (DEGENERATE).
 FT ZN_FING 485 507 C2H2-TYPE.
 FT ZN_FING 513 535 C2H2-TYPE.
 FT ZN_FING 541 563 C2H2-TYPE.
 SQ SEQUENCE 581 AA; 65644 MW; C78724098648C84F CRC64;

Query Match 90.0%; Score 36; DB 1; Length 581;
 Best Local Similarity 83.3%; Pred. No. 21;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YKGLC 6
 :|||||
 Db 370 FKGLC 375

RESULT 5

ID Z319_HUMAN STANDARD; PRT; 582 AA.
 AC Q9P2F9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 319.
 GN ZNF319 OR KIAA1388.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20181126; PubMed=10718198;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirotsawa M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XVI.
 RT The complete sequences of 150 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:65-73(2000).
 CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB037809; BAA92626.1; ALT_INIT.
 CC Genew; HGNC:13644; ZNF319.
 CC HSSP; P08045; 12NF.
 CC InterPro; IPR007087; Znf_C2H2.
 CC InterPro; IPR007086; Znf_C2H2_sub.
 CC Pfam; PF00096; zf-C2H2; 9.
 CC ProDom; PD000003; Znf_C2H2; 4.
 CC SMART; SM00355; Znf_C2H2; 14.
 CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
 CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 15.
 KW Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;
 FT DOMAIN 8 76 PRO-RTCH.
 FT ZN_FING 76 100 C2H2-TYPE.

FT ZN_FING 104 126 C2H2-TYPE (DEGENERATE).
 FT ZN_FING 132 154 C2H2-TYPE.
 FT ZN_FING 202 224 C2H2-TYPE.
 FT ZN_FING 230 252 C2H2-TYPE.
 FT ZN_FING 258 280 C2H2-TYPE (DEGENERATE).
 FT ZN_FING 287 309 C2H2-TYPE.
 FT ZN_FING 315 337 C2H2-TYPE.
 FT ZN_FING 343 365 C2H2-TYPE.
 FT ZN_FING 371 393 C2H2-TYPE.
 FT ZN_FING 399 421 C2H2-TYPE (DEGENERATE).
 FT ZN_FING 428 450 C2H2-TYPE.
 FT ZN_FING 458 480 C2H2-TYPE (DEGENERATE).
 FT ZN_FING 486 508 C2H2-TYPE.
 FT ZN_FING 514 536 C2H2-TYPE.
 FT ZN_FING 542 564 C2H2-TYPE.
 SQ SEQUENCE 582 AA; 65546 MW; E2FBD02256BE0A67 CRC64;

Query Match 90.0%; Score 36; DB 1; Length 582;
 Best Local Similarity 83.3%; Pred. No. 21;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 YKGLC 6
 :|||||
 Db 371 FKGLC 376

RESULT 6

ID ZN45_HUMAN STANDARD; PRT; 582 AA.
 AC Q02386; P78472;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 45 (BRC1744).
 GN ZNF45.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97220056; PubMed=9067431;
 RA Constantinou-Deltas C.D., Bashlades E., Patsalis P.C., Lee J.E.;
 RA Hadjinarrou M., Kroisel P.M., Ioannou P.A., Roses A.D., Lee J.E.;
 RT "Complete coding sequence, exon/intron arrangement and chromosome
 RT location of ZNF45, a KRAB-domain-containing gene.";
 RL Cytogenet. Cell Genet. 75:230-233(1996).
 CC [2]
 CC SEQUENCE OF 1-400 FROM N.A.
 CC MEDLINE=92217982; PubMed=1559709;
 RA Constantinou-Deltas C.D., Gilbert J., Bartlett R.J., Herbstreith M.,
 RA Roses A.D., Lee J.E.;
 RT "The identification and characterization of KRAB-domain-containing
 RT zinc finger proteins.";
 RL Genomics 12:581-589(1992).
 CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -----
 CC -!- SIMILARITY: Contains 1 KRAB domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L75847; AAB05653.1; -;
 CC EMBL; M67509; AAA36133.1; -;
 CC EMBL; M67509; AAA36134.1; -;
 CC HSSP; P08046; 1A11.

DR TRANSFAC; T04988; -
 DR Genew; HGNC:13111; ZNF45.
 DR MIM; 601781; -
 DR GO; GO:0003700; F:transcription factor activity; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.
 DR InterPro; IPR007067; Znf_C2H2.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00096; zf-C2H2; 15.
 DR ProDom; PD000003; Znf_C2H2; 13.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf_C2H2; 15.
 DR PROSITE; PS50805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 15.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 18.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat; Polymorphism.
 FT DOMAIN 8 78
 FT ZN_FING 192 214 C2H2-TYPE.
 FT ZN_FING 276 298 C2H2-TYPE.
 FT ZN_FING 304 326 C2H2-TYPE.
 FT ZN_FING 332 354 C2H2-TYPE.
 FT ZN_FING 360 382 C2H2-TYPE.
 FT ZN_FING 388 410 C2H2-TYPE.
 FT ZN_FING 416 438 C2H2-TYPE.
 FT ZN_FING 444 466 C2H2-TYPE.
 FT ZN_FING 472 494 C2H2-TYPE.
 FT ZN_FING 500 522 C2H2-TYPE.
 FT ZN_FING 528 550 C2H2-TYPE.
 FT ZN_FING 556 578 C2H2-TYPE.
 FT ZN_FING 584 606 C2H2-TYPE.
 FT ZN_FING 612 634 C2H2-TYPE.
 FT ZN_FING 640 662 C2H2-TYPE.
 FT VARIANT 187 187 A->T (IN dbSNP:1047452).
 FT VARIANT 255 255 /FTID-VAR_012019.
 FT VARIANT 299 299 R->K (IN dbSNP:399098).
 FT VARIANT 299 299 /FTID-VAR_012020.
 FT VARIANT 303 303 T->A (IN dbSNP:388706).
 FT VARIANT 303 303 /FTID-VAR_012021.
 FT VARIANT 504 504 P->R (IN dbSNP:388685).
 FT VARIANT 504 504 /FTID-VAR_012022.
 FT VARIANT 504 504 R->K (IN dbSNP:407731).
 FT VARIANT 504 504 /FTID-VAR_012023.
 FT SEQUENCE 682 AA; 78242 MW; 9C8E53E800FCE8B CRC64;
 Query Match 87.5%; Score 35; DB 1; Length 682;
 Best Local Similarity 83.3%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 YKCGLC 6
 Db 472 YKCGTC 477
 RESULT 7
 Z219_HUMAN STANDARD; PRT; 722 AA.
 AC Q9P2Y4; Q9BW28;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 219.
 GN ZNF219.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE-Testis;
 RX MEDLINE-20277481; PubMed-10819330;
 RA Sakai T., Toyoda A., Hashimoto K., Maeda H.;

RT *Isolation and characterization of a novel zinc finger gene, ZNF219,
 and mapping to the human chromosome 14q11 region.*;
 RN DNA Res. 7:137-141(2000).
 RP SEQUENCE FROM N.A.
 RC TISSUE-Kidney;
 RX MEDLINE-22388257; PubMed-12477932;
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT *Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.*;
 RL PROC. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB015427; BAA90526.1; -;
 DR EMBL; BC000694; AAH00694.1; -;
 DR Genew; HGNC:13011; ZNF219.
 DR MIM; 605036; -;
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0003700; F:transcription factor activity; TAS.
 DR InterPro; IPR007087; znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 8.
 DR SMART; SM00355; Znf_C2H2; 9.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat.
 FT ZN_FING 57 79 C2H2-TYPE.
 FT ZN_FING 85 107 C2H2-TYPE.
 FT ZN_FING 163 186 C2H2-TYPE.
 FT ZN_FING 274 296 C2H2-TYPE.
 FT ZN_FING 302 324 C2H2-TYPE.
 FT ZN_FING 498 520 C2H2-TYPE.
 FT ZN_FING 232 233 MISSING (IN REF. 2).
 FT CONFLICT 232 233 O -> E (IN REF. 2).
 FT CONFLICT 436 436 O -> E (IN REF. 2).
 SQ SEQUENCE 722 AA; 76876 MW; B19DA77B148BC45B CRC64;
 Query Match 87.5%; Score 35; DB 1; Length 722;
 Best Local Similarity 83.3%; Pred. No. 38;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YKCGLC 6
 Db 189 WKCGLC 194

RESULT 8

YAA4_SCHPO STANDARD; PRT; 1115 AA.
 AC Q09798;
 DT 01-NOV-1995 (Rel. 32, Last Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C22G7.04 in chromosome I.
 GN SPAC22G7.04.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 ON NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicakert G., Aert R., Robben J., Grymoprez B.,
 RA Weidjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RL The genome sequence of Schizosaccharomycetes pombe.*;
 RL Nature 413:871-880(2002).
 CC -!- SIMILARITY: TO YEAST PAN2 AND TO C.ELEGANS F31E3.4.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; Z54328; CAA91128.1; -
 DR PIR; T11614;
 DR GeneDB_Spombe; SPAC22G7.04; -
 DR InterPro; IPR006055; Exonuclease.
 DR InterPro; IPR001394; UCH-2.
 DR Pfam; PF00929; Exonuclease; 1.
 DR Pfam; PF00443; UCH; 1.
 DR SMART; SM00479; EXOIII; 1.
 DR SMART; SM00320; WD40; 2.
 DR PROSITE; PS50235; UCH_2_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 1115 AA; 126919 MW; EEF6F15E1897269D CRC64;

Query Match 87.5%; Score 35; DB 1; Length 1115;
 Best Local Similarity 66.7%; Pred. No. 56;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6

Db 640 YKCGIC 645

|||||

RESULT 9
 OVOL_HUMAN
 ID OVOL_HUMAN STANDARD; PRT; 235 AA.
 AC O14753;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Putative transcription factor Ovo-like 1 (hovol1) (Fragment).
 GN OVOL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98051941; PubMed=9383297;
 RA Chidambaram A., Allikmets R., Chandrasekarappa S., Guru S.C., Modi W.,
 RA Gerrard B., Dean M.;
 RT Characterization of a human homolog (OVOL1) of the Drosophila ovo
 RT gene, which maps to chromosome 11q13.*;
 RL Mamm. Genome 8:950-951(1997).
 CC -!- FUNCTION: PUTATIVE TRANSCRIPTION FACTOR. INVOLVED IN HAIR
 CC FORMATION AND SPERMATOGENESIS. MAY FUNCTION IN THE DIFFERENTIATION
 CC AND/OR MAINTENANCE OF THE UROGENITAL SYSTEM (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity). AND ALSO IN ADULT
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN FETAL KIDNEY, AND ALSO IN ADULT
 CC PANCREAS AND PLACENTA. NOT EXPRESSED IN INTESTINE, PERIPHERAL
 CC BLOOD LYMPHOCYTES AND OVARY.
 CC -!- SIMILARITY: Contains 4 C2H2-type zinc fingers.
 CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; TWO FRAMESHIFTS WERE
 CC CORRECTED TO EXTEND THE SEQUENCE IN THE N-TERMINAL SO AS TO
 CC MAXIMIZE THE SIMILARITY WITH THE MOUSE ORTHOLOG. BUT IT WAS NOT
 CC POSSIBLE TO RECOVER THE FULL SEQUENCE.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 CC FMBL; AF016045; AAB72084.1; ALT_FRAME.
 DR HSSP; P15822; IBBO.
 DR Genew; HGNC:8525; OVOL1.
 DR MIM; 602313; -
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003700; P:transcription factor activity; NAS.
 DR GO; GO:0008355; P:regulation of transcription, DNA-dependent; NAS.
 DR GO; GO:0007283; P:spermatogenesis; NAS.
 DR InterPro; IPR007087; Znf.C2H2.
 DR Pfam; PF00096; zf-C2H2; 4.
 DR ProDom; PD0000003; Znf.C2H2; 1.
 DR SMART; SM00355; Znf.C2H2; 4.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
 KW Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat;
 KW Transcription regulation.
 FT NON_TER 1
 FT ZN_FING 86 108 C2H2-TYPE 1.
 FT ZN_FING 114 136 C2H2-TYPE 2.
 FT ZN_FING 142 165 C2H2-TYPE 3.
 FT ZN_FING 181 203 C2H2-TYPE 4.
 SQ SEQUENCE 235 AA; 26388 MW; 511C8A5A7CDB6FE CRC64;

Query Match 85.0%; Score 34; DB 1; Length 235;
 Best Local Similarity 83.3%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 YKGLC 6
DB      142 YKSLC 147

RESULT 10
OVOL_MOUSE
ID      OVOL_MOUSE      STANDARD;      PRT;      267 AA.
AC      Q9WTJ2;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DE      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Putative transcription factor Ovo-like 1 (mOvol) (mOvola).
GN      OVOL1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Tissue-Skin;
RA      Schonbaum C.P., Fantes J., Mahowald A.P.;
RT      "Characterization of mouse and Caenorhabditis elegans genes related to
RT      the Drosophila melanogaster ovo/svb gene.";
RL      Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      FUNCTION.
RX      MEDLINE=99026118; PubMed=9808631;
RA      Dai X., Schonbaum C., Degenstein L., Bai W., Mahowald A., Fuchs E.;
RT      "The ovo gene required for cuticle formation and oogenesis in flies is
RT      involved in hair formation and spermatogenesis in mice.";
RL      Genes Dev. 12:3452-3463(1998).
CC      -1- FUNCTION: PUTATIVE TRANSCRIPTION FACTOR. INVOLVED IN HAIR
CC      FORMATION AND SPERMATOGENESIS. MAY FUNCTION IN THE DIFFERENTIATION
CC      AND/OR MAINTENANCE OF THE UROGENITAL SYSTEM.
CC      -1- SUBCELLULAR LOCATION: Nuclear.
CC      -1- TISSUE SPECIFICITY: EXPRESSED IN SKIN, TESTIS, KIDNEY AND WEAKLY
CC      IN LUNG. NOT DETECTED IN HEART, BRAIN, SPLEEN, LIVER AND SKELETAL
CC      MUSCLE.
CC      -1- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT E14.5 DAY IN THE
CC      SUPRABASAL LAYERS OF DEVELOPING EPIDERMIS, AT E15.5 EXPRESSION
CC      BEGINS IN THE INNER CELLS OF DEVELOPING HAIR GERMS AND RESTRICTED
CC      TO INNER ROOT SHEATH AND/OR PRECORTICAL CELLS OF DEVELOPING HAIR
CC      FOLLICLES.
CC      -1- SIMILARITY: Contains 4 C2H2-type zinc fingers.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF134804; AAD29689.1; -.
DR      EMBL; AF134805; AAD29690.1; -.
DR      HSP; P07248; 2ADR.
DR      MGD; MGI:1330290; Ovoll.
DR      GO; GO:0005634; C:nucleus; IDA.
DR      GO; GO:0008544; P:epidermal differentiation; IMP.
DR      GO; GO:0007498; P:mesoderm development; IMP.
DR      GO; GO:0007283; P:spermatogenesis; IMP.
DR      InterPro; IPR007087; Znf.C2H2.
DR      Pfam; PF00096; zf-C2H2; 4.
DR      ProDom; PD000003; Znf.C2H2; 1.
DR      SMART; SM00355; Znf.C2H2; 4.
DR      PROSITE; PS00028; ZINC_FINGER_C2H2.1; 3.
DR      PROSITE; PS00157; ZINC_FINGER_C2H2.2; 3.
KW      Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat;
Transcription regulation.
ET      ZN_FING 116 140 C2H2-TYPE 1.
FT      ZN_FING 146 168 C2H2-TYPE 2.
FT      ZN_FING 174 197 C2H2-TYPE 3.

FT      ZN_FING 213 236 C2H2-TYPE 4.
SQ      SEQUENCE 267 AA; 30221 MW; DAD4F51150C21C2D CRC64;

Query Match      85.0%; Score 34; DB 1; Length 267;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 YKGLC 6
DB      174 YKSLC 179

RESULT 11
Z185_MOUSE
ID      Z185_MOUSE      STANDARD;      PRT;      352 AA.
AC      Q62394;
DT      15-JUL-1999 (Rel. 38, Created)
DT      15-JUL-1999 (Rel. 38, Last sequence update)
DE      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Zinc finger protein 185 (LIM-domain protein ZFP185) (P1-A).
GN      ZNF185 OR ZFP185.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=96425694; PubMed=8828036;
RA      Levin M.L., Chatterjee A., Pradiola A., Worley K.C., Wennert M.,
RA      Zhuchenko O., Smith R.F., Lee C.C., Herman G.E.;
RT      "A comparative transcription map of the murine bare patches (Bpa) and
RT      striated (Str) critical regions and human Xq28.";
RL      Genome Res. 6:465-477(1996).
CC      -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF CELLULAR
CC      PROLIFERATION AND/OR DIFFERENTIATION (BY SIMILARITY).
CC      -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC      -1- TISSUE SPECIFICITY: EXPRESSED IN SKIN, KIDNEY, OVARY, TESTIS. ALSO
CC      EXPRESSED IN BRAIN, CARTILAGE, HEART, LUNG, SPLEEN AND THYMUS.
CC      -1- DEVELOPMENTAL STAGE: AT E14.5, ONLY EXPRESSED IN MESENCHYMAL
CC      CELLS. AT E16.5 EXPRESSED ALSO IN CELLS LINING THE VERTEBRAE AND
CC      TENDONS OF THE PROXIMAL TAIL. IN LATE EMBRYOGENESIS, EXPRESSED IN
CC      MESENCHYMAL CELLS ADJACENT TO THE DISTAL LIMB BONES (TIBIA AND
CC      CANCANEUM), IN TENDONS AND IN THE CONNECTIVE TISSUE SHEATH
CC      (EPIMYSIUM) SURROUNDING THE SKELETAL MUSCLE. ALSO EXPRESSED IN THE
CC      EPITHELIA OF THE EPIDIDYMIS OF THE TESTIS.
CC      -1- SIMILARITY: Contains 1 LIM zinc-binding domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U46687; AAC52628.1; -.
DR      MGD; MGI:108095; Zfp185.
DR      InterPro; IPR001781; LIM.
DR      ProDom; PD000094; LIM; 1.
DR      SMART; SM00132; LIM; 1.
DR      PROSITE; PS00478; LIM_DOMAIN_1; FALSE_NEG.
DR      PROSITE; PS0023; LIM_DOMAIN_2; 1.
KW      LIM domain; Metal-binding; Zinc.
FT      DOMAIN 23 26 POLY-GLUJ.
FT      DOMAIN 292 347 LIM.
SQ      SEQUENCE 352 AA; 38322 MW; 2AB1F833D7AF1A5C CRC64;

Query Match      85.0%; Score 34; DB 1; Length 352;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 YKGLC 6
          :|||:|

```

Query Match 85.0%; Score 34; DB 1; Length 472;
Best Local Similarity 83.3%; Pred. No. 40;

Matches	5;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	YKGLC 6							
Db	335	YKALC 340							
RESULT 14									
MLP2.DROME									
ID	MLP2.DROME	STANDARD:	PRT:	495	AA.				
AC	Q24400; Q9VI62;								
DT	01-NOV-1997 (Rel. 35, Created)								
DT	01-NOV-1997 (Rel. 35, Last sequence update)								
DT	16-OCT-2001 (Rel. 40, Last annotation update)								
DE	Muscle LIM protein MLP84B.								
GN	LIM3 OR MLP84B OR CG10699.								
OS	Drosophila melanogaster (Fruit fly).								
OC	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;								
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;								
OC	Ephydroidea; Drosophilidae; Drosophila.								
OX	NCBI_TaxID=7227;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	MEDLINE=96387325; PubMed=8794860;								
RA	Stronach B.E., Siegrist S.E., Beckerle M.C.;								
RT	"Two muscle-specific LIM proteins in Drosophila.";								
RL	J. Cell Biol. 134:1179-1195(1996).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RA	Stronach B.E., Renfranz P.J., Lilly B., Beckerle M.C.;								
RT	"Muscle LIM proteins associate with muscle sarcomeres and require								
RL	dMEF2 for their expression during Drosophila myogenesis.";								
RN	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=Berkeley;								
RX	MEDLINE=20196006; PubMed=10731132;								
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,								
RA	Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,								
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,								
RA	Sutton R.C., Rogers Y.-H.C., Blasej K.G., Champagne M., Pfeiffer B.D.,								
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,								
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,								
RA	Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,								
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,								
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,								
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,								
RA	Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,								
RA	De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,								
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,								
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,								
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,								
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,								
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,								
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,								
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,								
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,								
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,								
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,								
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,								
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,								
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,								
RA	Palazzolo M., Pittman G.S., Pan S., Poilard J., Puri V., Reese M.G.,								
RA	Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,								
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,								
RA	Spiel E., Spradling A.C., Stapleton M., Strong R., Sun E.,								
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,								
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,								
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,								
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,								
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,								
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;								

```

RT *The genome sequence of Drosophila melanogaster.*;
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: PLAYS A ROLE IN CELL DIFFERENTIATION LATE IN MYOGENESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSION IS RESTRICTED TO THE
CC SOMATIC, VICERAL, AND PHARYNGEAL MUSCLES, WITHIN THE SOMATIC
CC MUSCULATURE. MLP48B IS LOCALIZED AT THE ENDS OF MUSCLES FIBERS AT
CC THE POINT OF ATTACHMENT TO THE EPIDERMIS. THERE IS NO EXPRESSION
CC IN CARDIAC MESODERM OR IN FAT BODY.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS BIPHASIC, PEAKING LATE IN
CC EMBRYOGENESIS (16-24 H EMBRYOS) AND DURING THE LARVAL TO PUPAL
CC TRANSITION, WHEN THE MUSCULATURE IS DIFFERENTIATING. FOUND IN
CC DEVELOPING MUSCLES OF THE VISCERAL AND SOMATIC MESODERM SUBSEQUENT
CC TO THE FORMATION OF THE MUSCLE PRECURSOR CELLS. DECREASED LEVELS
CC ARE STILL DETECTABLE IN ADULTS.
CC -1- SIMILARITY: TO THE VERTEBRATE CYSTEINE-RICH PROTEINS.
CC -1- SIMILARITY: Contains 5 LIM zinc-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X91245; CAA62627.1; -.
DR EMBL; AF090832; AAC61591.1; -.
DR EMBL; AF003672; AAF54063.1; -.
DR HSSP; P32965; ICTL.
DR FlyBase; FBgn0014863; Mlp84B.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 5.
DR ProDom; PD000094; LIM; 5.
DR SMART; SM00132; LIM; 5.
DR PROSITE; PS00478; LIM_DOMAIN_1; 5.
DR PROSITE; PSS0023; LIM_DOMAIN_2; 5.
KW Nuclear protein; Repeat; LIM domain; Metal-binding; Zinc; Myogenesis;
KW Developmental protein; Differentiation.
FT DOMAIN 12 63
FT DOMAIN 65 80
FT DOMAIN 66 71
FT DOMAIN 120 172
FT DOMAIN 175 180
FT DOMAIN 178 189
FT DOMAIN 222 274
FT DOMAIN 226 291
FT DOMAIN 325 377
FT DOMAIN 379 390
FT DOMAIN 421 473
FT DOMAIN 475 490
SO SEQUENCE 495 AA; 53525 MW; 2E559B9178E54C0E CRC64;

Query Match      85.0%; Score 34; DB 1; Length 495;
Best Local Similarity 66.7%; Pred.No. 41;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YKGLC 6
Db 37 FKCGMC 42

RESULT 15
Y619.METJA
ID Y619.METJA STANDARD; PRT; 506 AA.
AC Q58036;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0619.
GN MJ0619.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;

```

```

OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A. 2661 / ATCC 43067;
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=9637999; PubMed=868087;
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.
RL Science 273:1058-1073(1996).
CC -I- SIMILARITY: SOME, TO B.SUBTILIS NARA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U67510; AAB98614.1; -.
DR PIR; C64377; C64377.
DR TIGR; MJ0619; -.
DR InterPro; IPR006638; E1p3.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; E1p3; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 506 AA; 57481 MW; 25FA1F6A380398D8 CRC64;
SQ
Query Match 85.0%; Score 34; DB 1; Length 506;
Best Local Similarity 83.3%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 YKCGLC 6
Db 75 YDCGIC 80

```

Search completed: October 1, 2003, 10:42:41
Job time : 8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:34:26 ; Search time 13.1667 seconds
(without alignments)
43.824 Million cell updates/sec

Title: US-10-057-890A-6

Perfect score: 40

Sequence: 1 YKCGLC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	427	2 G81260	UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diamino-pimelate
2	40	100.0	614	2 JH0500	zinc finger protein
3	40	100.0	1350	2 S00647	finger protein - A
4	37	92.5	57	2 B43284	zinc finger protein
5	37	92.5	149	2 A95105	conserved hypother
6	37	92.5	149	2 A97973	conserved hypother
7	36	90.0	1638	2 T25352	hypothetical prote
8	35	87.5	582	2 C90165	conserved hypother
9	35	87.5	1000	2 S74069	hypothetical prote
10	35	87.5	1115	2 T11614	probable poly(A)-s
11	34	85.0	186	2 T22278	hypothetical prote
12	34	85.0	190	2 T22279	hypothetical prote
13	34	85.0	277	2 T27824	hypothetical prote
14	34	85.0	309	2 F83044	nitrate-inducible
15	34	85.0	396	2 C84797	hypothetical prote
16	34	85.0	472	2 T34222	transcription fact
17	34	85.0	506	2 C68377	hypothetical prote
18	34	85.0	610	2 G72494	hypothetical prote
19	34	85.0	621	2 T22904	hypothetical prote
20	34	85.0	1028	2 A56038	DNA-binding protei
21	34	85.0	1186	2 T33754	O/E-1-associated z
22	34	85.0	1213	2 S16356	ovo protein - frui
23	34	85.0	1323	2 T24534	regulatory protein
24	33	82.5	138	2 T21792	hypothetical prote
25	33	82.5	201	2 I57505	zinc finger protei
26	33	82.5	223	2 F31201	GLI-related finger
27	33	82.5	225	2 D69090	conserved hypother
28	33	82.5	233	2 D84797	hypothetical prote
29	33	82.5	240	2 S34619	probable fumarate

30	33	82.5	265	2 T33695	hypothetical prote
31	33	82.5	265	2 T15486	hypothetical prote
32	33	82.5	283	2 G96979	uncharacterized pr
33	33	82.5	293	2 T22385	hypothetical prote
34	33	82.5	319	2 T46469	hypothetical prote
35	33	82.5	395	2 D70441	oxido/reductase 1r
36	33	82.5	395	2 T52423	C2H2 zinc finger p
37	33	82.5	427	2 A35659	krueppel-related p
38	33	82.5	501	1 F64462	polyferredoxin 2 -
39	33	82.5	544	2 T27444	hypothetical prote
40	33	82.5	555	2 I53869	zinc finger protei
41	33	82.5	577	2 B37057	integrin beta-6 ch
42	33	82.5	580	2 A46538	Ig heavy chain, se
43	33	82.5	626	2 F72256	re-hydrogenase, su
44	33	82.5	656	2 JC2005	integrin beta-5 ch
45	33	82.5	693	2 I37570	zinc finger protei

ALIGNMENTS

RESULT 1

G81260

UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diamino-pimelate ligase (EC 6.3.2.13) Cj
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: G81260

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Ch

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; B

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveal

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: G81260

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-427 <PAR>

A:Cross-references: GB:AL139079; GB:AL111168; MID:g6968971; PIDN:CAB73629.1; PID:g6

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: murE; Cj1641

C:Superfamily: UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase

C:Keywords: ligase

Query Match 100.0%; Score 40; DB 2; Length 427;
Best Local Similarity 100.0%; Pred.No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6

DB 86 YKCGLC 91

RESULT 2

JH0500

zinc finger protein 29 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Dec-2000

C:Accession: JH0500

R:Denny, P.; Ashworth, A.

Gene 106, 221-227, 1991

A:Title: A zinc finger protein-encoding gene expressed in the post-meiotic phase of

A:Reference number: JH0500; MUID:92039080; PMID:1937051

A:Accession: JH0500

A:Molecule type: mRNA

A:Residues: 1-614 <DEN>

A:Cross-references: GB:X55126; MID:g55470; PIDN:CAA38920.1; PID:g55471

A:Experimental source: testis

C:Keywords: DNA binding; zinc finger

F:217-244/Region: zinc finger

F:245-272/Region: zinc finger

F:273-300/Region: zinc finger

F:301-328/Region: zinc finger

F:329-356/Region: zinc finger

F:357-384/Region: zinc finger
 F:385-412/Region: zinc finger
 F:413-440/Region: zinc finger
 F:441-468/Region: zinc finger
 F:469-496/Region: zinc finger
 F:497-524/Region: zinc finger
 F:525-552/Region: zinc finger
 F:553-580/Region: zinc finger
 F:581-608/Region: zinc finger

Query Match 100.0%; Score 40; DB 2; Length 614;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 |||||
 Db 446 YKCGLC 451

RESULT 3

S00647
 finger protein - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 31-Dec-1993
 C:Accession: S00647
 B:Ruiz i Altaba, A.; Perry-O'Keefe, H.; Melton, D.A.
 EMBO J. 6, 3065-3070, 1987
 A:Title: Xfin: an embryonic gene encoding a multifingered protein in Xenopus.
 A:Reference number: S00647; MUID:88082679; PMID:2826129
 A:Accession: S00647
 A:Molecule type: mRNA
 A:Residues: 1-1350 <RUI>
 A:Cross-references: ENBL:X06021
 A:Note: it is uncertain whether Met-1, Met-11, Met-16, Met-38, or Met-39 is the initiator
 C:Genetics:
 A:Gene: fin
 C:Keywords: DNA binding; zinc finger

Query Match 100.0%; Score 40; DB 2; Length 1350;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 |||||
 Db 1044 YKCGLC 1049

RESULT 4

B43284
 zinc finger protein ZNF48 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
 C:Accession: B43284
 R:Lichter, P.; Bray, P.; Ried, T.; Dawid, I.B.; Ward, D.C.
 Genomics 13, 995-1007, 1992
 A:Title: Clustering of C2-H2 zinc finger motif sequences within telomeric and fragile sites
 A:Reference number: A43284; MUID:92372070; PMID:1505991
 A:Accession: B43284
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-57 <LIC>
 A:Cross-references: GB:M88358; NID:g340451; PIDN:AAA61316.1; PID:g340452
 A:Note: sequence extracted from NCBI backbone (NCBIN:111632, NCBI:P:111633)

Query Match 92.5%; Score 37; DB 2; Length 57;
 Best Local Similarity 83.3%; Pred. No. 17;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 |||||
 Db 15 YKCGVC 20

RESULT 5

A95105
 conserved hypothetical protein SP0909 [imported] - Streptococcus pneumoniae (strain conserved)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: A95105
 R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; On, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappeler, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morri
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: A95105

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-149 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK75034.1; PID:g14972383; GSPDB:GN00164; TIGR

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0909

Query Match 92.5%; Score 37; DB 2; Length 149;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 |||||
 Db 132 YKCGLC 137

RESULT 6

A97973
 conserved hypothetical protein spr0809 [imported] - Streptococcus pneumoniae (strain conserved)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C:Accession: A97973
 R:Hoskins, J.A.; Arnold, J.; Blaszcak, L.; Bargett, S.; DeHoff, B.S
 e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S
 Y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S
 A:Title: Genome of the Bacterium Streptococcus pneumoniae strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:1154234

A:Accession: A97973

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-149 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAK99613.1; PID:g15458409; GSPDB:GN00174

C:Genetics:

A:Gene: spr0809

Query Match 92.5%; Score 37; DB 2; Length 149;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 |||||
 Db 132 YKCGLC 137

RESULT 7

T25352

hypothetical protein T27C5.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T25352

R:Cummings, P.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z20020

A:Accession: T25352

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A;Residues: 1-1638 <WIL>
A;Cross-references: EMBL:282058; PIDN:CAB04870.1; GSPDB:GN00023; CESP:T27C5.9
A;Experimental source: clone T27C5
C;Genetics:
A;Gene: CESP:T27C5.9
A;Map position: 5
A;Introns: 1432/2

Query Match 90.0%; Score 36; DB 2; Length 1638;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
I:|||||
Db 159 YKCGLC 164

RESULT 8
C90165
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C;Accession: C90165
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayer, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90165
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-582 <KUR>
A;Cross-references: GB:AE006641; NID:gl13813374; PIDN:AAK40578.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS00235
C;Superfamily: conserved hypothetical protein MTH831

Query Match 87.5%; Score 35; DB 2; Length 582;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKCGLC 6
I:|||||
Db 122 YKCGLC 127

RESULT 9
S74069
hypothetical protein c0138 - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 17-Mar-1999
C;Accession: S74069
R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A;Title: Organizational characteristics and information content of an archaeal genome: 1
A;Reference number: S73076; MUID:97055432; PMID:8899719
A;Accession: S74069
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1000 <SEN>
A;Cross-references: EMBL:Y08256; NID:gl1707679; PID:e283937; PID:gl1707768
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996

Query Match 87.5%; Score 35; DB 2; Length 1000;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
I:|||||
Db 573 HKCGLC 578

RESULT 10

T11614

probable poly(A)-specific ribonuclease (EC 3.1.13.4) - fission yeast (Schizosacchar-
C;Species: Schizosaccharomyces pombe
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 29-Oct-1999
C;Accession: T11614; S62448
R;Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z17300
A;Accession: T11614
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1115 <BAR>

A;Cross-references: EMBL:254328; NID:gl009451; PID:gl009454

A;Experimental source: strain 972h(-)

R;Badcock, K.; Churcher, C.M.

submitted to the EMBL Data Library, October 1995

A;Reference number: S62445

A;Accession: S62448

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1115 <BAD>

A;Cross-references: EMBL:254328; NID:gl009451; PIDN:CAA91128.1; PID:gl009454

C;Genetics:

A;Map position: IL

A;Introns: 6/3; 39/3; 95/2; 164/3; 357/2; 793/1

A;Note: SPAC2267.04

C;Keywords: hydrolase

Query Match 87.5%; Score 35; DB 2; Length 1115;

Best Local Similarity 66.7%; Pred. No. 2.4e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6

I:|||||

Db 640 YKCGIC 645

RESULT 11

T22278

hypothetical protein F46B3.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000

C;Accession: T22278

R;Ainscough, R.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19541

A;Accession: T22278

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-186 <WIL>

A;Cross-references: EMBL:281540; PIDN:CAB04402.1; GSPDB:GN00023; CESP:F46B3.1

A;Experimental source: clone F46B3

C;Genetics:

A;Gene: CESP:F46B3.1

A;Map position: 5

A;Introns: 42/2; 79/2; 115/2; 155/2

C;Superfamily: Caenorhabditis elegans hypothetical protein F49F1.6

Query Match 85.0%; Score 34; DB 2; Length 186;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKCGLC 6

I:|||||

Db 46 YKCGLC 51

RESULT 12

T22279

hypothetical protein F46B3.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000

C:Accession: T22279

R:Ainscough, R.

Submitted to the EMBL Data Library, November 1996

A:Reference number: Z19541

A:Accession: T22279

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-190 <WIL>

A:Cross-references: EMBL:Z81540; PIDN:CA804403.1; GSPDB:GN00023; CESP:F46B3.2

A:Experimental source: clone F46B3

C:Genetics:

A:Gene: CESP:F46B3.2

A:Map position: 5

A:Introns: 94/2; 119/2; 159/2

C:Superfamily: Caenorhabditis elegans hypothetical protein F49F1.6

Query Match 85.0%; Score 34; DB 2; Length 190;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKGLC 6

DB 61 YTCGLC 66

RESULT 13

T27824

hypothetical protein ZK287.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27824

R:McMurray, A.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z20425

A:Accession: T27824

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-277 <WIL>

A:Cross-references: EMBL:Z70757; PIDN:CAA94802.1; GSPDB:GN00023; CESP:ZK287.6

A:Experimental source: clone ZK287

C:Genetics:

A:Gene: CESP:ZK287.6

A:Map position: 5

A:Introns: 37/2; 66/3; 144/1

Query Match 85.0%; Score 34; DB 2; Length 277;

Best Local Similarity 66.7%; Pred. No. 1.4e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6

DB 206 YECGIC 211

RESULT 14

F83044

nitrate-inducible formate dehydrogenase, beta subunit PA4811 [Imported] - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 02-Aug-2002

C:Accession: F83044

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: F83044

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-309 <STO>

A:Cross-references: GB:AE004894; GB:AE004091; NID:99951076; PIDN:AAG08197.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Superfamily: formate dehydrogenase, nitrate-inducible, beta chain; ferredoxin 2[4F

Query Match 85.0%; Score 34; DB 2; Length 309;

Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKGLC 6

DB 154 YKCSLC 159

RESULT 15

C84797

hypothetical protein At2g37800 [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84797

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vente

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: C84797

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-396 <STO>

A:Cross-references: GB:AE002093; NID:94895204; PIDN:AAD32791.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g37800

A:Map position: 2

Query Match

Best Local Similarity 85.0%; Score 34; DB 2; Length 396;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKGLC 6

DB 71 YKCGQC 76

Search completed: October 1, 2003, 10:47:51

Job time: 14.1667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:38:06 ; Search time 12.5 Seconds
(without alignments)
20.309 Million cell updates/sec

Title: US-10-057-890A-6
Perfect score: 40
Sequence: 1 YKGLC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	25	3	US-09-253-396A-216
2	40	100.0	25	6	Sequence 216, App
3	40	100.0	26	2	Patent No. 5196510
4	35	87.5	127	2	Sequence 55, Appl
5	34	85.0	27	2	Sequence 4659, Ap
6	34	85.0	29	3	Sequence 76, Appl
7	34	85.0	310	4	Sequence 217, App
8	33	82.5	352	4	Sequence 27339, A
9	33	82.5	378	4	Sequence 118, App
10	33	82.5	398	4	Sequence 122, App
11	33	82.5	399	4	Sequence 121, App
12	33	82.5	488	2	Sequence 18242, A
13	33	82.5	488	2	Sequence 17, Appl
14	33	82.5	577	2	Sequence 17, Appl
15	33	82.5	577	2	Sequence 29, Appl
16	33	82.5	577	4	Sequence 29, Appl
17	33	82.5	799	1	Sequence 29, Appl
18	32	80.0	26	2	Sequence 38, Appl
19	32	80.0	28	2	Sequence 45, Appl
20	32	80.0	29	3	Sequence 45, Appl
21	32	80.0	29	2	Sequence 4, Appl
22	32	80.0	95	4	Sequence 7, Appl
23	32	80.0	224	2	Sequence 2, Appl
24	32	80.0	402	4	Sequence 5, Appl
25	32	80.0	409	4	Sequence 2, Appl
26	32	80.0	676	2	Sequence 40, Appl
27	32	80.0	676	3	Sequence 40, Appl

28 32 80.0 976 3 US-08-894-997-50 Sequence 50, Appl
29 32 80.0 1332 2 US-08-971-244-2 Sequence 2, Appl
30 32 80.0 1332 2 US-09-286-891-2 Sequence 2, Appl
31 31 77.5 26 2 US-08-620-151-53 Sequence 53, Appl
32 31 77.5 41 4 US-09-389-831-15 Sequence 15, Appl
33 31 77.5 43 1 US-08-208-108-7 Sequence 7, Appl
34 31 77.5 46 3 US-08-955-937A-10 Sequence 10, Appl
35 31 77.5 46 3 US-09-300-985-10 Sequence 10, Appl
36 31 77.5 84 3 US-09-386-493-8 Sequence 8, Appl
37 31 77.5 84 4 US-09-996-243-246 Sequence 246, App
38 31 77.5 105 4 US-09-389-956-70 Sequence 70, Appl
39 31 77.5 109 3 US-08-961-564A-9 Sequence 9, Appl
40 31 77.5 119 4 US-09-107-532A-6580 Sequence 6580, Ap
41 31 77.5 181 4 US-09-702-114A-15 Sequence 15, Appl
42 31 77.5 224 3 US-09-045-193-2 Sequence 2, Appl
43 31 77.5 227 2 US-08-471-717-2 Sequence 82, Appl
44 31 77.5 344 4 US-08-311-731A-82 Sequence 4, Appl
45 31 77.5 419 2 US-08-574-959A-4

ALIGNMENTS

RESULT 1
US-09-253-396A-216
; Sequence 216, Application US/09253396A
; Patent No. 6205404
; GENERAL INFORMATION:
; APPLICANT: Genome Dynamics, Inc.
; TITLE OF INVENTION: DNA-Binding Proteins of the Zinc-Finger Class
; FILE REFERENCE: 1116242-0003 file: genome03F.app
; CURRENT APPLICATION NUMBER: US/09/253,396A
; CURRENT FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 216
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-253-396A-216

Query Match 100.0%; Score 40; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
DB 1 YKGLC 6

RESULT 2
5196510-11
; Patent No. 5196510
; APPLICANT: RODRIGUEZ, JOHN D.; MCKEARN, THOMAS J.; ALVAREZ,
; VERNON, L.; RADCLIFFE, ROBERT D.
; TITLE OF INVENTION: MOLECULAR RECOGNITION UNITS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/519,702
; FILING DATE: 07-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 291,730
; FILING DATE: 29-DEC-1988
; SEQ ID NO:11:
; LENGTH: 25
5196510-11

Query Match 100.0%; Score 40; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YKGLC 6
|||||

Db 1 YKCGLC 6

RESULT 3
US-08-620-151-55
; Sequence 55, Application US/08620151
; Patent No. 5928955
; GENERAL INFORMATION:
; APPLICANT: Imperiali, Barbara
; APPLICANT: Walkup, Grant K.
; TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR
; TITLE OF INVENTION: DIVALENT ZINC
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,151
; FILING DATE: 22-MAR-1996
; CLASSIFICATION: 422
; ATTORNEY/AGENT INFORMATION:
; NAME: Shannon, Karen L.
; REGISTRATION NUMBER: 36,675
; REFERENCE/DOCKET NUMBER: 8597/6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-620-151-55

Query Match 100.0%; Score 40; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
Db 2 YKCGLC 7

RESULT 4
US-09-107-532A-4659
; Sequence 4659, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4659:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...127
; SEQUENCE DESCRIPTION: SEQ ID NO: 4659:
US-09-107-532A-4659

Query Match 87.5%; Score 35; DB 4; Length 127;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
Db 7 YRCGIC 12

RESULT 5
US-08-620-151-76
; Sequence 76, Application US/08620151
; Patent No. 5928955
; GENERAL INFORMATION:
; APPLICANT: Imperiali, Barbara
; APPLICANT: Walkup, Grant K.
; TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR
; TITLE OF INVENTION: DIVALENT ZINC
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,151
; FILING DATE: 22-MAR-1996
; CLASSIFICATION: 422
; ATTORNEY/AGENT INFORMATION:
; NAME: Shannon, Karen L.
; REGISTRATION NUMBER: 36,675
; REFERENCE/DOCKET NUMBER: 8597/6

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-620-151-76

Query Match 85.0%; Score 34; DB 2; Length 27;
Best Local Similarity 83.3%; Pred. No. 9.4;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKGLC 6
DB 2 YPCGLC 7

RESULT 6
US-09-253-396A-217
Sequence 217, Application US/09253396A
Patent No. 6205404
GENERAL INFORMATION:
APPLICANT: Genome Dynamics, Inc.
TITLE OF INVENTION: DNA-Binding Proteins of the Zinc-Finger Class
FILE REFERENCE: 116242-0003 file: genome03f.app
CURRENT APPLICATION NUMBER: US/09/253.396A
CURRENT FILING DATE: 1999-02-19
NUMBER OF SEQ ID NOS: 231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 217
LENGTH: 29
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-253-396A-217

Query Match 85.0%; Score 34; DB 3; Length 29;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKGLC 6
DB 6 YPCGLC 11

RESULT 7
US-09-252-991A-27339
Sequence 27339, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27339
LENGTH: 310
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27339

Query Match 85.0%; Score 34; DB 4; Length 310;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKGLC 6
DB 155 YKSLC 160

RESULT 8

US-08-979-847B-118
Sequence 118, Application US/08979847B
Patent No. 6582703
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
BESEME, FREDERIC
BEDIN, FREDERIC
PARANHOS-BACCALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TUKU, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PRE
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/979,847B
FILING DATE: 26-No. 6582703-1997

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 118:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 118:
US-08-979-847B-118

Query Match 82.5%; Score 33; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KGLC 6
DB 85 KGLC 89

RESULT 9

US-08-979-847B-122
Sequence 122, Application US/08979847B
Patent No. 6582703
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
BESEME, FREDERIC

BEVIN, FREDERIC
PARANHOS-BACCALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TUKÉ, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847B
FILING DATE: 26-No. 6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 122:
US-08-979-847B-122
Query Match 82.5%; Score 33; DB 4; Length 378;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 KCGLC 6
DB 98 KCGLC 102
RESULT 10
US-08-979-847B-121
Sequence 121, Application US/08979847B
Patent No. 6582703
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
BESEME, FREDERIC
BEVIN, FREDERIC
PARANHOS-BACCALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TUKÉ, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC

STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847B
FILING DATE: 26-No. 6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 121:
US-08-979-847B-121
Query Match 82.5%; Score 33; DB 4; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 KCGLC 6
DB 118 KCGLC 122
RESULT 11
US-09-252-991A-18242
Sequence 18242, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18242
LENGTH: 399
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18242
Query Match 82.5%; Score 33; DB 4; Length 399;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 KCGLC 6
DB 358 FRCGLC 363
RESULT 12
US-08-933-750C-17

```
; Sequence 17, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933-750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-853-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT03
; CLONE: 641127
; US-08-933-750C-17

Query Match 82.5%; Score 33; DB 2; Length 488;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKCGLC 6
Db 347 YKCGEC 352

RESULT 13
US-09-234-613-17
; Sequence 17, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
```

```
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-853-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT03
; CLONE: 641127
; US-09-234-613-17

Query Match 82.5%; Score 33; DB 3; Length 488;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKCGLC 6
Db 347 YKCGEC 352

RESULT 14
US-07-728-215-29
; Sequence 29, Application US/07728215
; Patent No. 5962643
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Dean
; APPLICANT: Quaranta, Vito
; APPLICANT: Pytela, Robert
; TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Prettly Schroeder, Brueggemann & Clark
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 577 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-728-215-29

Query Match 82.5%; Score 33; DB 2; Length 577;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YKCGLC 6
Db 333 YQCGVC 338
|:|:|:|

RESULT 15
US-08-938-085A-29
Sequence 29, Application US/08938085A
Patent No 6339148
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pytela, Robert
TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3634
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,085A
FILING DATE: 26-SEP-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,215
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parent, Annette S.
REGISTRATION NUMBER: 42,058
REFERENCE/DOCKET NUMBER: 023070-080210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 577 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-938-085A-29

Query Match 82.5%; Score 33; DB 4; Length 577;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YKCGLC 6
Db 333 YQCGVC 338
|:|:|:|

Search completed: October 1, 2003, 10:49:12
Job time : 13.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:04:56 ; Search time 38.3333 Seconds
(without alignments)
24.844 Million cell updates/sec

Title: US-10-057-890A-6
Perfect score: 40
Sequence: 1 YKCGLC 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

Database : A_Geneseq_19Jun03.*

1:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query #		Length	DB	ID	Description
		Match	Mismatch				
1	40	100.0	6	23	ABG32535	Scaffold domain pe	
2	40	100.0	25	12	AAR15297	Metal binding effe	
3	40	100.0	138	23	ABG32539	Human CCR5-based s	
4	40	100.0	157	23	ABG32540	Human CCR5-based s	
5	38	95.0	357	22	ABG32635	Human protein sequ	
6	38	95.0	632	22	AAU705085	Novel human diagno	
7	38	95.0	632	22	ABG19011	Novel human diagno	
8	38	95.0	769	24	AAE33782	Human nucleic acid	
9	37	92.5	139	21	AAG04173	Arabidopsis thalia	

10	37	92.5	139	21	AAG39378	Arabidopsis thaliana
11	37	92.5	149	24	ABU01311	S. pneumoniae type
12	37	92.5	193	21	AGC04172	Arabidopsis thaliana
13	37	92.5	193	21	AAG39377	Arabidopsis thaliana
14	37	92.5	197	21	AAG04171	Arabidopsis thaliana
15	37	92.5	538	22	ABB71212	Drosophila melanog
16	37	92.5	614	23	ABG95656	Human nucleic acid
17	37	92.5	645	22	ABG28885	Novel human diago
18	37	92.5	702	23	ABBO4333	Human zinc finger
19	37	92.5	799	21	ABE21003	Human nucleic acid
20	36	90.0	359	21	AAB34522	Human secreted pro
21	36	90.0	360	21	AAB34521	Human secreted pro
22	36	90.0	361	21	AAB34470	Human secreted pro
23	36	90.0	361	22	AAM78800	Human protein SEQ
24	36	90.0	393	22	AM79784	Human protein SEQ
25	36	90.0	582	23	ABB97429	Novel human protei
26	36	90.0	751	22	ABG25861	Novel human diago
27	36	90.0	3067	22	ABG69362	Drosophila melanog
28	35	87.5	250	23	ABF64010	Human ORF380. Hom
29	35	87.5	302	24	ABU07872	Human polypeptide
30	35	87.5	312	22	ABU07872	Human zinc finger
31	35	87.5	313	22	AAU90653	Human secreted pro
32	35	87.5	316	22	AAU16014	Human novel secret
33	35	87.5	316	24	ABU55083	Human novel polype
34	35	87.5	402	22	ABU50185	Human transcriptio
35	35	87.5	622	22	ABB59242	Drosophila melanog
36	35	87.5	722	22	AM39029	Human polypeptide
37	34	85.0	723	22	AAU08511	zinc finger domain
38	37	92.5	23	24	ABR40929	Human zinc finger
39	34	85.0	107	22	AAU07367	Human polypeptide
40	34	85.0	117	22	AAAG4023	Human polypeptide
41	34	85.0	158	23	ABG60027	Human DTHP polype
42	34	85.0	186	22	ABG65567	Drosophila melanog
43	34	85.0	201	22	AAU30581	Novel human secret
44	34	85.0	233	23	ABG60046	Human DTHP polype
45	34	85.0	233	24	ABU05313	Human diagnostics

ALIGNMENTS

RESULT 1
ABG32535
ID ABG32535 standard; peptide; 6 AA.

Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection; ECD; extracellular domain; metal chelating motif; zinc finger protein; integral membrane protein; soluble loop; intracellular domain; ICD; gene therapy; immunogen; viral infection; scaffold domain; human.

CC involved in HIV (human immunodeficiency virus) infection).

CC Also included are: (1) a polypeptide comprising a scaffold domain; (2) a

CC nucleic acid encoding the fusion polypeptide; (3) a vector cassette for

CC the expression of the fusion polypeptide comprising an expression region

CC operably linked to a promoter, where the expression region comprises a

CC number of cassettes, each of which encodes a module, domain or strand of

CC the fusion polypeptide and (4) a host cell comprising the vector or

CC nucleic acid. The fusion polypeptide is useful for screening molecules

CC that bind/activate/inhibit/modulate the fusion polypeptide, by expressing

CC the fusion polypeptide from and identifying a molecule that binds to the

CC fusion polypeptide. The fusion polypeptide is useful in diagnostic

CC methods, in assays to identify compounds that interact with loops of

CC fragments of an extracellular domain (ECD) or an intracellular domain

CC (ICD) or to rapidly assay the function of mutated portions of mutant

CC integral membrane proteins without having to produce significant

CC quantities of the entire mutant integral membrane proteins from which they

CC are designed, to competitively bind the ligand of a naturally occurring

CC receptor in vitro or in vivo, to display and/or screen soluble domains

CC from protein such as integral membrane proteins, to probe the structure

CC of ECD or ICD, or both, of an integral protein membrane, to modulate the

CC activity of a receptor in vivo, and for treating or preventing viral

CC infection, preferably human HIV infection e.g. by gene therapy using the

CC encoding nucleic acid. The present sequence is a scaffolded protein

CC based on the ECD region of human CCR5 (not defined).

XX SQ Sequence 138 AA;

Query Match 100.0%; Score 40; DB 23; Length 138;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6

DB 31 YKCGLC 36

RESULT 4

ABG32540

ID ABG32540 standard; protein; 157 AA.

AC ABG32540;

DT 15-NOV-2002 (first entry)

DE Human CCR5-based scaffolded fusion protein #2.

KW Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;

KW ECD; extracellular domain; metal chelating motif; zinc finger protein;

KW integral membrane protein; soluble loop; intracellular domain; ICD;

KW gene therapy; immunogen; viral infection; human.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

XX FT Peptide 1..19

XX FT Protein 20..157

XX FT /label= Signal_peptide

XX FT /label= Mature_scaffolded_protein

PN WO200260477-A1.

XX 08-AUG-2002.

XX 29-JAN-2002; 2002WO-US02377.

XX 31-JAN-2001; 2001US-265782P.

XX 31-JAN-2001; 2001US-265858P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Coleman TA, Mansfield B;

XX WPI; 2002-643357/69.

XX Novel scaffolded fusion polypeptide useful for therapeutic purposes or

XX for screening molecules that bind/activate/inhibit/modulate the

XX polypeptide, comprises a functional polypeptide domain fused to a

XX scaffold domain -

XX Example 2; Page 41; 64pp; English.

XX The invention relates to a scaffolded fusion polypeptide comprising a

XX functional polypeptide domain fused to a scaffold domain, where the

XX integral polypeptide domain corresponds to a soluble loop of an

XX integral membrane protein (e.g. human CCR5, a transmembrane receptor

XX involved in HIV (human immunodeficiency virus) infection).

XX Also included are: (1) a polypeptide comprising a scaffold domain; (2) a

XX nucleic acid encoding the fusion polypeptide; (3) a vector cassette for

XX the expression of the fusion polypeptide comprising an expression region

XX operably linked to a promoter, where the expression region comprises a

XX number of cassettes, each of which encodes a module, domain or strand of

XX the fusion polypeptide and (4) a host cell comprising the vector or

XX nucleic acid. The fusion polypeptide is useful for screening molecules

XX that bind/activate/inhibit/modulate the fusion polypeptide, by expressing

XX the fusion polypeptide from and identifying a molecule that binds to the

XX fusion polypeptide. The fusion polypeptide is useful in diagnostic

XX methods, in assays to identify compounds that interact with loops of

XX fragments of an extracellular domain (ECD) or an intracellular domain

XX (ICD) or to rapidly assay the function of mutated portions of mutant

XX integral membrane proteins without having to produce significant

XX quantities of the entire mutant integral membrane proteins from which they

XX are designed, to competitively bind the ligand of a naturally occurring

XX receptor in vitro or in vivo, to display and/or screen soluble domains

XX from protein such as integral membrane proteins, to probe the structure

XX of ECD or ICD, or both, of an integral protein membrane, to modulate the

XX activity of a receptor in vivo, and for treating or preventing viral

XX infection, preferably human HIV infection e.g. by gene therapy using the

XX encoding nucleic acid. The present sequence is a scaffolded protein

XX based on the ECD region of human CCR5 (not defined).

XX SQ Sequence 157 AA;

Query Match 100.0%; Score 40; DB 23; Length 157;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6

DB 50 YKCGLC 55

RESULT 5

AAB93635

ID AAB93635 standard; Protein; 357 AA.

XX AAB93635;

DT 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:13123.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

XX Claim 8; SEQ ID 13123; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AA03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AA092446 to
 CC AA095893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 357 AA;

Query Match 95.0%; Score 38; DB 22; Length 357;

Best Local Similarity 83.3%; Pred. No. 92;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 |||||
 Db 334 YKCGIC 339

RESULT 6

ID AAU30585 standard; Protein; 620 AA.

XX AAU30585;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #1076.

XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 PT Claim 20; Page 314; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

XX Sequence 620 AA;

Query Match 95.0%; Score 38; DB 22; Length 620;

Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 |||||
 Db 359 YKCGIC 364

RESULT 7

ID ABG19011 standard; Protein; 632 AA.

XX ABG19011;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #19002.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB: AAS83198.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID No 49370; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 632 AA:
SQ
Query Match 95.0%; Score 38; DB 22; Length 632;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 YKCGLC 6
| | | | |
Db 609 YKCGIC 614

RESULT 8
AAE33782
ID AAE33782 standard; Protein; 769 AA.
XX
XX AAE33782;
XX
XX 16-APR-2003 (first entry)
XX
XX Human nucleic acid associated protein (NAAP)-22.
XX
XX Human; nucleic acid associated protein; NAAP; cancer; atherosclerosis;
KW cell proliferative disorder; neurological disorder; Huntington's disease;
KW epilepsy; stroke; immune disorder; acquired immune deficiency syndrome;
KW AIDS; inflammatory disorder; allergy; developmental disorder; infection;
KW hypothyroidism; cushing's syndrome; gene therapy; cytostatic; nootropic;
KW anticonvulsant; neuroprotective; cerebroprotective; thyromimetic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 24..85
FT Domain /note= "Krab box domain"
FT Domain 26..97
FT Domain /note= "Krab box domain"
FT Domain 413..428
FT Domain /note= "Transmembrane domain"
FT Domain 160..180
FT Domain /note= "Zinc finger C2H2 type domain"
FT Domain 188..208
FT Domain /note= "Zinc finger C2H2 type domain"
FT Domain 216..236

FT Domain /note= "Zinc finger C2H2 type domain"
FT 244..264
FT Domain /note= "Zinc finger C2H2 type domain"
FT 272..292
FT Domain /note= "Zinc finger C2H2 type domain"
FT 300..320
FT Domain /note= "Zinc finger C2H2 type domain"
FT 356..376
FT Domain /note= "Zinc finger C2H2 type domain"
FT 384..404
FT Domain /note= "Zinc finger C2H2 type domain"
FT 412..432
FT Domain /note= "Zinc finger C2H2 type domain"
FT 440..460
FT Domain /note= "Zinc finger C2H2 type domain"
FT 488..488
FT Domain /note= "Zinc finger C2H2 type domain"
FT 496..516
FT Domain /note= "Zinc finger C2H2 type domain"
FT 524..544
FT Domain /note= "Zinc finger C2H2 type domain"
FT 552..572
FT Domain /note= "Zinc finger C2H2 type domain"
FT 580..600
FT Domain /note= "Zinc finger C2H2 type domain"
FT 608..628
FT Domain /note= "Zinc finger C2H2 type domain"
FT 636..656
FT Domain /note= "Zinc finger C2H2 type domain"
FT 664..684
FT Domain /note= "Zinc finger C2H2 type domain"
FT 692..712
FT Domain /note= "Zinc finger C2H2 type domain"
FT 720..740
FT Domain /note= "Zinc finger C2H2 type domain"
FT 748..768
FT Domain /note= "Zinc finger C2H2 type domain"
XX
XX WO200299115-A2.
XX
XX 12-DEC-2002.
XX
XX 31-MAY-2002; 2002WO-US17050.
XX
XX 01-JUN-2001; 2001US-295359P.
XX 08-JUN-2001; 2001US-296878P.
XX 08-JUN-2001; 2001US-297222P.
XX 15-JUN-2001; 2001US-298615P.
XX 15-JUN-2001; 2001US-298665P.
XX 15-JUN-2001; 2001US-298693P.
XX 21-JUN-2001; 2001US-300176P.
XX 19-APR-2002; 2002US-373891P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX (YDEH/) YUE H.
XX
XX Yue H, Tang YT, Baughn MR, Becha SD, Warren BA, Wallia NK, Lal PG;
PI Lee EA, Hafalia AJA, Richardson TW, Griffin JA, Emerling BM;
PI Ramkumar J, Yue H, Swarnakar A, Tran B, Li JX, Yao MG, Yang J;
PI Ison CH, Forsythe IJ, Honnell CD, Arvizu CS, Elliott VS, Lu Y;
PI Ding L, Luo W, Wang YE, Burford N, Borowsky ML, Nguyen DB;
PI Chinn AM, Kable AE;
XX WPI; 2003-140626/13.
XX N-PSDB; AAD51696.
XX
XX New human nucleic acid associated proteins (NAAP), useful for
PT diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis,
XX epilepsy, or infections -
XX
XX Claim 78; Column 214-215; 257pp; English.
XX

CC The present invention relates to human nucleic acid associated proteins
 CC (NAAP) and polynucleotides encoding such proteins. NAAP sequences are
 CC useful in diagnosing, treating and preventing diseases or conditions
 CC associated with the decreased expression or overexpression of NAAP such
 CC as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome)
 CC disorders or infections. They are also used in gene therapy. The present
 CC sequence is human NAAP-22 protein.
 XX
 SQ Sequence 769 AA;

Query Match 95.0%; Score 38; DB 24; Length 769;

Best Local Similarity 83.3%; Pred. No. 1.8e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGIC 6

||||:|

Db 746 YKCGIC 751

RESULT 9

ARG04173

ID AAG04173 standard; Protein: 139 AA.

AC AAG04173;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 99.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127452.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142190.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 92.5%; Score 37; DB 21; Length 139;
Best Local Similarity 83.3%; Pred. No. 59;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
|||||
Db 21 YKGVG 26

RESULT 10

AAG39378
ID AAG39378 standard; Protein: 139 AA.

XX AC AAG39378;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 48713.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144089.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 21 YKCGVC 26

RESULT 11

ABU01311
 ID ABU01311 standard; Protein; 149 AA.

AC ABU01311;

DT 11-FEB-2003 (first entry)

DE S. pneumoniae type 4 strain protein from coding region #886.

KW Bacterial meningitis; pneumonia; sepsis; otitis media;
 KW ear infection; antiinflammatory; antibacterial; immunostimulant;
 KW auditory; respiratory; gene therapy; vaccine.

OS Streptococcus pneumoniae type 4 strain.

PN WO200277021-A2.

PD 03-OCT-2002.

PF 27-MAR-2002; 2002WO-IB02163.

XX 27-MAR-2001; 2001GB-0007658.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Masignani V, Tettelin H, Fraser C;

XX WPI: 2003-040579/03.

DR N-PSDB; ABX06598.

XX New proteins and nucleic acid molecules from Streptococcus pneumoniae.
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
 PT or ear infection

PS Claim 1; SEQ ID NO 1772; 56pp; English.

XX The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS5454. Also included are an antibody which binds one of the
 CC proteins, treating a patient by administering the protein, DNA or
 CC antibody (in a composition), a kit comprising first and second primers,
 CC which are the nucleic acid cited above or fragments between nucleotides
 CC 8-100 of a sequence not defined in the specification, for amplifying a
 CC target sequence contained within a Streptococcus nucleic acid sequence,
 CC where the first primer is substantially complementary to the target
 CC sequence and the second primer is substantially complementary to the
 CC complement of the target sequence, and where the parts of the primers
 CC having substantial complementarity define the termini of the target
 CC sequence to be amplified, assay comprising contacting a test compound
 CC with the protein, and determining whether the test compound binds to the
 CC protein and a Streptococcus pneumoniae bacterium, where one or more
 CC genes encoding the proteins has been rendered inactive. The proteins,
 CC nucleic acid molecules, antibody and compositions are useful as
 CC medicaments for treating or preventing a disease or infection due to
 CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
 CC sepsis, otitis media or ear infection. They are also useful in developing
 CC vaccines, diagnostics and antibiotics. The methods are useful for
 CC identifying immunodominant proteins. The present sequence is one of
 CC the 2469 proteins expressed by the identified coding regions from the
 CC genomic sequence.

CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 149 AA;

Query Match 92.5%; Score 37; DB 24; Length 149;
 Best Local Similarity 83.3%; Pred. No. 63;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKCGLC 6
 DB 132 YKCGLC 137

RESULT 12

AAG04172
 ID AAG04172 standard; Protein; 193 AA.

XX AAG04172;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 98.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139113.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147030.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match          92.5%; Score 37; DB 21; Length 193;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 YKCGLC 6
Db       75 YKCGVC 80

RESULT 13
AAG39377
ID AAG39377 standard; Protein; 193 AA.
AC AAG39377;
XX
XX
DT 18-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48712.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 18-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

```

```
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      92.5%; Score 37; DB 21; Length 193;
Best Local Similarity 83.3%; Pred.No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKGLC 6
DB 75 YKCGVC 80
```

RESULT 14

AAG04171

ID AAG04171 standard; Protein; 197 AA.

XX

AC AAG04171;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 97.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 06-APR-1999; 99US-0128234.

PR

PR 08-APR-1999; 99US-0128714.

PR

PR 16-APR-1999; 99US-0129845.

PR

PR 19-APR-1999; 99US-0130077.

PR

PR 21-APR-1999; 99US-0130449.

PR

PR 23-APR-1999; 99US-0130510.

PR

PR 23-APR-1999; 99US-0130891.

PR

PR 28-APR-1999; 99US-0131449.

PR

PR 30-APR-1999; 99US-0132048.

PR

PR 30-APR-1999; 99US-0132407.

PR

PR 04-MAY-1999; 99US-0132484.

PR

PR 05-MAY-1999; 99US-0132485.

PR

PR 06-MAY-1999; 99US-0132486.

PR

PR 06-MAY-1999; 99US-0132487.

PR

PR 07-MAY-1999; 99US-0132863.

PR

PR 11-MAY-1999; 99US-0134256.

PR

PR 14-MAY-1999; 99US-0134218.

PR

PR 14-MAY-1999; 99US-0134219.

PR

PR 14-MAY-1999; 99US-0134221.

PR

PR 18-MAY-1999; 99US-0134370.

PR

PR 18-MAY-1999; 99US-0134768.

PR

PR 19-MAY-1999; 99US-0134941.

PR

PR 20-MAY-1999; 99US-0135124.

PR

PR 21-MAY-1999; 99US-0135353.

PR

PR 24-MAY-1999; 99US-0135629.

PR

PR 25-MAY-1999; 99US-0136021.

PR

PR 27-MAY-1999; 99US-0136392.

PR

PR 28-MAY-1999; 99US-0136782.

PR

PR 01-JUN-1999; 99US-0137222.

PR

PR 03-JUN-1999; 99US-0137528.

PR

PR 04-JUN-1999; 99US-0137502.

PR

PR 07-JUN-1999; 99US-0137724.

PR

PR 08-JUN-1999; 99US-0138094.

PR

PR 10-JUN-1999; 99US-0138540.

PR

PR 10-JUN-1999; 99US-0138847.

PR

PR 14-JUN-1999; 99US-0139119.

PR

PR 16-JUN-1999; 99US-0139452.

PR

PR 16-JUN-1999; 99US-0139453.

PR

PR 17-JUN-1999; 99US-0139492.

PR

PR 18-JUN-1999; 99US-0139454.

PR

PR 18-JUN-1999; 99US-0139455.

PR

PR 18-JUN-1999; 99US-0139456.

PR

PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 21-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149375.
 PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 26-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 92.5%; Score 37; DB 21; Length 197;
 Best Local Similarity 83.3%; Pred. NO. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YKGLC 6
 |||||
 Db 79 YKGCVC 84

RESULT 15
 ABB71212
 ID ABB71212 standard: Protein: 538 AA.
 XX
 AC ABB71212;

```

XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 40428.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX N-PSDB; ABL15315.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 40428; 2lpp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 538 AA;
Query Match 92.5%; Score 37; DB 22; Length 538;
Best Local Similarity 83.3%; Pred. NO. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 YKCGLC 6
Db 491 YKCGVC 496
Search completed: October 1, 2003, 10:41:52
Job time : 40.3333 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:04:56 ; Search time 31.9444 Seconds
(without alignments)
24.844 Million cell updates/sec

Title: US-10-057-890A-7

Perfect score: 30

Sequence: 1 HORVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	5	23	ABG32536 Scaffold domain pe
2	30	100.0	14	22	AAW97437 Human peptide #712
3	30	100.0	15	23	AAW51635 Human zinc finger
4	30	100.0	15	23	AAW51639 Human actin 15 N-t
5	30	100.0	23	4	ABR41005 Human zinc finger
6	30	100.0	30	22	ABG54217 Human liver peptid
7	30	100.0	30	22	ABG39268 Peptide #6774 enco
8	30	100.0	30	22	AAW59932 Human brain expres
9	30	100.0	30	22	AAW72526 Human bone marrow

10	30	100.0	30	22	AAW32770 Peptide #6807 enco
11	30	100.0	30	23	ABG42346 Human peptide enco
12	30	100.0	51	22	AAU48414 Propionibacterium
13	30	100.0	58	21	AAU26880 Zea mays protein f
14	30	100.0	61	23	ABP08539 Human ORFX protein
15	30	100.0	70	22	AAU48510 Propionibacterium
16	30	100.0	88	22	AAU16535 Human novel secret
17	30	100.0	88	24	ABU55604 Human novel polype
18	30	100.0	93	22	AAU62204 Human gene 13-enco
19	30	100.0	93	23	ABG63557 Human albumin fusi
20	30	100.0	93	24	ABP57726 A. thaliana EIN6 z
21	30	100.0	114	22	AAU16419 Human novel secret
22	30	100.0	114	24	ABU53488 Human novel polype
23	30	100.0	115	21	AAU56977 Human prostate can
24	30	100.0	120	22	AAU57790 Propionibacterium
25	30	100.0	120	22	AAU16017 Human novel secret
26	30	100.0	120	24	ABU55086 Human novel polype
27	30	100.0	125	22	ABU10507 Human cDNA SEQ ID
28	30	100.0	125	23	ABP67094 Human polypeptide
29	30	100.0	125	23	ABG70325 Human MDR1 protein
30	30	100.0	126	22	ABG54412 Human liver peptid
31	30	100.0	126	22	ABB39422 Peptide #6928 enco
32	30	100.0	126	22	AAW60103 Human brain expres
33	30	100.0	126	22	AAW72714 Human bone marrow
34	30	100.0	126	22	AAW32945 Peptide #6982 enco
35	30	100.0	126	23	ABG42538 Human peptide enco
36	30	100.0	127	20	AAU00358 Fragment of human
37	30	100.0	135	22	AAO05868 Human polypeptide
38	30	100.0	135	22	AAW5681 Human colon cancer
39	30	100.0	136	23	AAW51638 Human actin 15 pol
40	30	100.0	138	23	ABG32539 Human CCR5-based s
41	30	100.0	145	23	AAO14573 Novel murine zinc
42	30	100.0	148	21	ABW42669 Human ORFX ORF2433
43	30	100.0	149	22	AAU18479 human diagnostic a
44	30	100.0	152	21	AAU18887 Zea mays protein f
45	30	100.0	154	22	AAU16368 Human novel secret

ALIGNMENTS

RESULT 1
ABG32536

ID ABB32536 standard; peptide; 5 AA.

XX AC ABB32536;

XX DT 15-NOV-2002 (first entry)

XX DE Scaffold domain peptide #2.

XX KW Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;
KW ECD; extracellular domain; metal chelating motif; zinc finger protein;
KW integral membrane protein; soluble loop; intracellular domain; ICD;
KW gene therapy; immunogen; viral infection; scaffold domain; human.

XX OS Homo sapiens.

XX PN WO200260477-A1.

XX PD 08-AUG-2002.

XX PF 29-JAN-2002; 2002WO-US02377.

XX PR 31-JAN-2001; 2001US-265782P.

XX PR 31-JAN-2001; 2001US-265858P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Coleman TA, Mansfield B.

XX DR Peptide #6774 enco

XX WPI; 2002-643357/69.

PT Novel scaffolded fusion polypeptide useful for therapeutic purposes or
PT for screening molecules that bind/activate/inhibit/modulate the
PT polypeptide, comprises a functional polypeptide domain fused to a
XX scaffold domain .
PS Disclosure; Page 16; 64pp; English.
XX
XX The invention relates to a scaffolded fusion polypeptide comprising a
CC functional polypeptide domain fused to a scaffold domain, where the
CC functional polypeptide domain corresponds to a soluble loop of an
CC integral membrane protein (e.g. human CCR5, a transmembrane receptor
CC involved in HIV (human immunodeficiency virus) infection).
CC Also included are: (1) a polypeptide comprising a scaffold domain; (2) a
CC nucleic acid encoding the fusion polypeptide; (3) a vector cassette for
CC the expression of the fusion polypeptide comprising an expression region
CC operably linked to a promoter, where the expression region comprises a
CC number of cassettes, each of which encodes a module, domain or strand of
CC the fusion polypeptide and (4) a host cell comprising the vector or
CC nucleic acid. The fusion polypeptide is useful for screening molecules
CC that bind/activate/inhibit/modulate the fusion polypeptide, by expressing
CC the fusion polypeptide from and identifying a molecule that binds to the
CC fusion polypeptide. The fusion polypeptide is useful in diagnostic
CC methods, in assays to identify compounds that interact with loops of
CC fragments of an extracellular domain (ECD) or an intracellular domain
CC (ICD) or to rapidly assay the function of mutated portions of mutant
CC integral membrane proteins without having to produce significant
CC quantities of the entire mutant integral membrane protein, to generate
CC antibodies that recognise the integral membrane proteins from which they
CC are designed, to competitively bind the ligand of a naturally occurring
CC receptor in vitro or in vivo, to display and/or screen soluble domains
CC from protein such as integral membrane proteins, to probe the structure
CC of ECD or ICD, or both, of an integral protein membrane, to modulate the
CC activity of a receptor in vivo, and for treating or preventing viral
CC infection, preferably human HIV infection e.g. by gene therapy using the
CC encoding nucleic acid. The present sequence is a scaffold domain
CC suitable for inclusion in the fusion protein of the invention.
XX
XX Sequence 5 AA;
SQ
Query Match 100.0%; Score 30; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HORVH 5
DB 1 HORVH 5
|||||
RESULT 2
AAM97437
ID AAM97437 standard; Peptide: 14 AA.
XX
XX AAM97437;
AC
XX
XX 24-JAN-2002 (first entry)
DT
XX
XX Human peptide #712 encoded by a SNP oligonucleotide.
DE
XX
XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
XX Homo sapiens.
OS
XX
XX WO200147944-A2.
PN
XX
XX 05-JUL-2001.
PD
XX

PF 28-DEC-2000; 2000MO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shimkets RA, Leach M;
PI
XX WPI: 2001-465210/50.
DR
XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Disclosure; Page 3823; 4143pp; English.
XX
XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The oligonucleotides and the peptides encoded by one such oligonucleotide.
CC The present sequence is a peptide encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms.
XX
XX Sequence 14 AA;
SQ
Query Match 100.0%; Score 30; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HORVH 5
DB 8 HORVH 12
|||||
RESULT 3
AAM51635
ID AAM51635 standard; Peptide: 15 AA.
XX
XX AAM51635;
AC
XX
XX 18-FEB-2002 (first entry)
DT
XX
XX Human zinc finger protein 22 N-terminal peptide.
DE
XX
XX Human; zinc finger protein 22; cancer; nosohaemia; infection;
KW HIV; human immunodeficiency virus; immunological disease;
KW inflammatory disorder.
XX
XX Homo sapiens.
OS
XX
XX CN1307039-A.
PN
XX
XX 08-AUG-2001.
PD
XX
XX 26-JAN-2000; 2000CN-0111545.
PF
XX
XX 26-JAN-2000; 2000CN-0111545.
PR
XX (BODA-) BODAO GENE TECH CO LTD SHANGHAI.
PA
XX
XX Mao Y, Xie Y;
PI
XX

DR WPI: 2002-062741/09.
 XX Polypeptide-human zinc finger protein 22 and polynucleotide for coding
 PT said polypeptide -
 XX
 PS Example 6; Page 19 (disclosure): 34pp; Chinese.
 XX
 CC The invention relates to a novel polypeptide, human zinc finger protein
 CC 22. The polypeptide is useful for treating various diseases, such as
 CC malignant tumours, nosohaemia, HIV infection, immunological diseases and
 CC inflammatory disorder. The present sequence is the N-terminus of the
 CC polypeptide of the invention.
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 30; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HQRVH 5
 DB 3 HQRVH 7
 RESULT 4
 AAM51639
 ID AAM51639 standard; Peptide: 15 AA.
 AC AAM51639;
 XX
 DT 18-FEB-2002 (first entry)
 DE Human actin 15 N-terminal peptide.
 KW Human; zinc finger; PHD finger; actin 15; cancer; nosohaemia;
 KW HIV; human immunodeficiency virus; infection; immunological disease;
 KW inflammatory disorder.
 XX
 OS Homo sapiens.
 XX
 PN CN1307041-A.
 XX
 PD 08-AUG-2001.
 XX
 PF 26-JAN-2000; 2000CN-0111549.
 XX
 PR 26-JAN-2000; 2000CN-0111549.
 XX
 PA (BODA-) BODAO GENE TECH CO LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI: 2002-062743/09.
 XX
 PT Polypeptide-human zinc finger and PHD finger structural domain-contg
 PT actin 15 and polynucleotide for coding said polypeptide -
 XX
 PS Example 6; Page 20 (disclosure): 35pp; Chinese.
 XX
 CC The invention relates to a novel polypeptide, human zinc finger and
 CC PHD finger structural domain-contg actin 15. It is useful in the
 CC treatment of cancer, nosohaemia, HIV infection, immunological diseases
 CC and inflammatory disorders. The present sequence is the N-terminus
 CC of the polypeptide of the invention.
 XX
 SQ Sequence 15 AA;
 Query Match 100.0%; Score 30; DB 23; Length 15;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HQRVH 5
 DB 3 HQRVH 7

Db 2 HQRVH 6
 RESULT 5
 ABR41005
 ID ABR41005 standard; Protein; 23 AA.
 XX
 AC ABR41005;
 XX
 DT 21-MAY-2003 (first entry)
 XX
 DE Human zinc finger domain related protein sequence SEQ ID NO:181.
 XX
 KW Zinc finger domain; zinc finger; zinc finger binding domain;
 KW chimeric nucleic acid; library.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016571-A1.
 XX
 PD 27-FEB-2003.
 XX
 PF 17-AUG-2002; 2002WO-KR01560.
 XX
 PR 17-AUG-2001; 2001US-313402P.
 PR 22-APR-2002; 2002US-374355P.
 XX
 PA (TOOL-) TOOLGEN INC.
 XX
 PI Kim J, Bae K, Park K, Kwon Y, Ryu E, Hwang M;
 XX
 DR WPI: 2003-268344/26.
 DR N-PSDB; ACC41645.
 XX
 PT New library comprising polypeptides having zinc finger domains, useful
 PT for producing chimeric nucleic acids -
 XX
 PS Claim 3; Page 87; 234pp; English.
 XX
 CC The present invention describes a library comprising polypeptides. Each
 CC polypeptide comprises a first or second zinc finger domain. The domains
 CC of each polypeptide are identical to a zinc finger domain from a
 CC naturally occurring protein and either do not occur in the same naturally
 CC occurring protein or occur in the same naturally occurring protein in a
 CC different configuration than in the polypeptide. The domains vary among
 CC polypeptides. Also described: (1) producing chimeric nucleic acids;
 CC (2) generating an artificial zinc finger polypeptide that specifically
 CC binds to a target DNA site; and (3) identifying a nucleic acid encoding
 CC a zinc finger polypeptide that specifically recognises a target DNA site.
 CC The library can be used for producing chimeric nucleic acids, ACC41551.
 CC to ACC41758 and ABR40919 to ABR41015 represent nucleotide and amino acid
 CC sequences given in the exemplification of the present invention.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 30; DB 24; Length 23;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HQRVH 5
 DB 19 HQRVH 23
 RESULT 6
 ABG54217
 ID ABG54217 standard; Peptide; 30 AA.
 XX
 AC ABG54217;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID No 32865.

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX Homo sapiens.
 OS WO200157273-A2.
 PN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00664.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-48898/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human adult liver -
 XX Claim 27; SEQ ID NO 32865; 658pp; English.
 XX The invention relates to a single exon nucleic acid probe (SEN) (I) for
 XX measuring human gene expression in a sample derived from human adult
 XX liver, comprising one of 13109 defined nucleotide sequences given in the
 XX specification (or complements/ fragments). The probe hybridises at high
 XX stringency to a nucleic acid molecule expressed in the human adult
 XX liver. (I) may be used for predicting, measuring and displaying gene
 XX expression in samples derived from human adult liver. The genes
 XX identified may be involved in genetic liver diseases such as cirrhosis,
 XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 XX is associated with coronary heart disease. ABG47348-ABG59930 represent
 XX human liver single exon encoded peptides of the invention.
 XX Note: The sequence information for this patent does not appear in the
 XX printed specification but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 30 AA;
 Query Match 100.0%; Score 30; DB 22; Length 30;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HQRVH 5
 Db 21 HQRVH 25
 RESULT 7
 ID ABB39268 standard; Peptide; 30 AA.
 AC ABB39268;
 XX ABB39268;
 XX 04-FEB-2002 (first entry)
 DE Peptide #6774 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 OS WO200157277-A2.
 PN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00667.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human foetal liver -
 XX Claim 27; SEQ ID NO 31903; 639pp + sequence listing; English.
 XX The invention relates to a single exon nucleic acid probe for
 XX measuring human gene expression in a sample derived from human foetal
 XX liver. The single exon nucleic acid probes may be used for predicting,
 XX measuring and displaying gene expression in samples derived from human
 XX foetal liver. The present sequence is a peptide encoded by a single exon
 XX nucleic acid probe of the invention.
 XX Note: The sequence data for this patent did not form part of the
 XX printed specification, but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 30 AA;
 Query Match 100.0%; Score 30; DB 22; Length 30;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HQRVH 5
 Db 21 HQRVH 25
 RESULT 8
 ID AAM59932 standard; Protein; 30 AA.
 AC AAM59932;
 XX AAM59932;
 XX 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 32037.
 XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX Homo sapiens.
 OS WO200157275-A2.
 PN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00667.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

PD 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00669.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human foetal liver -
 XX Claim 27; SEQ ID NO 31903; 639pp + sequence listing; English.
 XX The invention relates to a single exon nucleic acid probe for
 XX measuring human gene expression in a sample derived from human foetal
 XX liver. The single exon nucleic acid probes may be used for predicting,
 XX measuring and displaying gene expression in samples derived from human
 XX foetal liver. The present sequence is a peptide encoded by a single exon
 XX nucleic acid probe of the invention.
 XX Note: The sequence data for this patent did not form part of the
 XX printed specification, but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 30 AA;
 Query Match 100.0%; Score 30; DB 22; Length 30;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HQRVH 5
 Db 21 HQRVH 25
 RESULT 8
 ID AAM59932 standard; Protein; 30 AA.
 AC AAM59932;
 XX AAM59932;
 XX 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 32037.
 XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX Homo sapiens.
 OS WO200157275-A2.
 PN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00667.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI: 2001-483446/52.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX Example 4: SEQ ID NO: 32037; 650pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 SQ Sequence 30 AA;
 Query Match 100.0%; Score 30; DB 22; Length 30;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HORVH 5
 DB 21 HORVH 25
 |||||
 RESULT 9
 AAM72526
 ID AAM72526 standard; Protein; 30 AA.
 XX
 AC AAM72526;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32832.
 XX
 DE Human bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 KW
 XX Homo sapiens.
 OS
 XX WO200157276-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00668.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI: 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX Example 4: SEQ ID NO: 32832; 658pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 SQ Sequence 30 AA;
 Query Match 100.0%; Score 30; DB 22; Length 30;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HORVH 5
 DB 21 HORVH 25
 |||||
 RESULT 10
 AAM32770
 ID AAM32770 standard; Protein; 30 AA.
 XX
 AC AAM32770;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #6807 encoded by probe for measuring placental gene expression.
 XX
 KW Probe: microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 OS
 XX Homo sapiens.
 XX WO200157272-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00663.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI: 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX Claim 27; SEQ ID NO 33039; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP:
 CC see AA131315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX
 SQ Sequence 30 AA;
 Query Match 100.0%; Score 30; DB 22; Length 30;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HORVH 5
 DB 21 HORVH 25
 |||||

RESULT 11

ABG42346
 ID ABG42346 standard; Peptide; 30 AA.
 AC ABG42346;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 32011.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00665.
 XX
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2002-114183/15.
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 XX Claim 27; SEQ ID No 32011; 634pp; English.
 XX
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene

CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.

CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 30 AA;

Query Match 100.0%; Score 30; DB 23; Length 30;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5

Db 21 HQRVH 25

RESULT 12

AAU48414
 ID AAU48414 standard; Protein; 51 AA.

XX AC AAU48414;

XX 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #9310.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59542.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

XX Example 1; SEQ ID No 9609; 1069pp; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 51 AA;

Query Match 100.0%; Score 30; DB 22; Length 51;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
 Db 31 HQRVH 35

RESULT 13

AAG26880
 ID AAG26880 standard; Protein: 58 AA.

XX AC AAG26880;

XX DT 17-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 31503.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hydriidisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

XX OS Zea mays subsp. mays.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132407.

XX PR 05-MAY-1999; 99US-0132484.

XX PR 06-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132487.

XX PR 11-MAY-1999; 99US-0132863.

XX PR 14-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145115.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.

CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORF proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 61 AA;
 Query Match 100.0%; Score 30; DB 23; Length 61;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HORVH 5
 Db | | | | |
 37 HORVH 41
 RESULT 15
 AAU48510
 ID AAU48510 standard: Protein: 70 AA.
 XX
 AC AAU48510;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #9406.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 9705; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 70 AA;

Query Match 100.0%; Score 30; DB 22; Length 70;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
 Db | | | | |
 49 HORVH 53

Search completed: October 1, 2003, 10:41:53
 Job time : 32.9444 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:34:26 ; Search time 10.9722 Seconds
(without alignments)
43.824 Million cell updates/sec

Title: US-10-057-890A-7

Perfect score: 30

Sequence: 1 HQRVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	30	100.0	39	2	D40984	finger protein zfa
2	30	100.0	56	2	S10398	finger protein kox
3	30	100.0	56	2	I37960	zinc finger protei
4	30	100.0	56	2	I37963	zinc finger protei
5	30	100.0	56	2	I37966	zinc finger protei
6	30	100.0	66	2	S47065	finger protein HZF
7	30	100.0	85	2	A42825	Kruppel-type zinc
8	30	100.0	87	2	I38941	zinc finger protei
9	30	100.0	88	2	D43284	finger protein ZNF
10	30	100.0	105	2	D48827	zinc finger protei
11	30	100.0	106	2	B31201	zinc finger protei
12	30	100.0	106	2	T34560	GLI-related finger
13	30	100.0	137	1	Q0BED9	hypothetical prote
14	30	100.0	139	2	I38620	HHLF3 protein - hu
15	30	100.0	169	2	A39240	zinc finger protei
16	30	100.0	172	2	S08575	finger protein mfg
17	30	100.0	183	2	S70007	finger protein (cl
18	30	100.0	194	2	I53859	finger protein zfo
19	30	100.0	196	2	S06558	zinc finger protei
20	30	100.0	201	2	I57505	zinc finger protei
21	30	100.0	209	2	S47068	finger protein HZF
22	30	100.0	216	2	S06781	finger protein (cl
23	30	100.0	223	2	S08576	finger protein (cl
24	30	100.0	223	2	S33951	finger protein ZNF
25	30	100.0	223	2	F31201	GLI-related finger
26	30	100.0	247	2	S06553	finger protein (cl
27	30	100.0	258	2	T34601	hypothetical prote
28	30	100.0	261	2	S70006	finger protein zfo
29	30	100.0	288	2	C56281	7alpha-cephem-meth

30	30	100.0	348	2	I38599	zinc finger protei
31	30	100.0	378	2	S33994	finger protein ZNF
32	30	100.0	386	2	T12527	hypothetical prote
33	30	100.0	393	2	JN0533	finger protein PML
34	30	100.0	401	2	A42177	KRAB-domain-contai
35	30	100.0	405	2	D70529	hypothetical prote
36	30	100.0	411	2	S10245	finger protein, te
37	30	100.0	420	2	S65084	finger protein XFG
38	30	100.0	421	2	AE2473	hypothetical prote
39	30	100.0	445	2	S32036	finger protein XFG
40	30	100.0	462	2	I51699	gene XGF 5.1c prot
41	30	100.0	475	2	S03679	finger protein (cl
42	30	100.0	482	2	S18210	hypothetical prote
43	30	100.0	485	2	A40751	finger protein MZF
44	30	100.0	518	1	FWOAG1	12S seed storage g
45	30	100.0	518	2	A36433	globulin precursor

ALIGNMENTS

RESULT 1

D40984
finger protein zfas8 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 08-Dec-2000
C:Accession: D40984
R:Crossley, P.H.; Little, P.F.R.
Proc. Natl. Acad. Sci. U.S.A. 88, 7923-7927, 1991
A:Title: A cluster of related zinc finger protein genes is deleted in the mouse emb

A:Reference number: A40984; MUID:91376058; PMID:1680234

A:Accession: D40984

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-39 <CRO>

A:Cross-references: GB:M74237

C:Keywords: DNA binding; zinc finger

Query Match 100.0%; Score 30; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 2.7; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0

Qy 1 HQRVH 5
Db 23 HQRVH 27

RESULT 2

S10398
finger protein kox10 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 21-Jul-2000
C:Accession: I37949; S10398
R:Thiesen, H.J.
New Biol. 2, 363-374, 1990

A:Title: Multiple genes encoding zinc finger domains are expressed in human T cells

A:Reference number: I37949; MUID:91145339; PMID:2288909

A:Accession: I37949

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-56 <RES>

A:Cross-references: EMBL:X52341; NID:g34142; PIDN:CAA36567.1; PID:g930072

C:Genetics:

A:Gene: GDB:ZNF17; HPF3

A:Cross-references: GDB:125324

A:Map position: 19q-19q

C:Keywords: DNA binding; zinc finger

Query Match 100.0%; Score 30; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.9; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0

Qy 1 HQRVH 5


```

|||||
Db      19  HORVH 23

RESULT 3
I37960
zinc finger protein kox20 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 05-Nov-1999
C:Accession: I37960; S10409
R:Thiesen, H.J.
New Biol. 2, 363-374, 1990
A:Title: Multiple genes encoding zinc finger domains are expressed in human T cells.
A:Reference number: I37949; MUID:91145339; PMID:2288909
A:Accession: I37960
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-56 <RES>
A:Cross-references: EMBL:X52351; NID:g34153; PIDN:CAA36577.1; PID:q930083
C:Genetics:
A:Gene: GDB:2NF26; K0X20
A:Cross-references: GDB:125333; OMIM:194537
A:Map position: 12q24.33-12q24.33
C:Keywords: DNA binding; zinc finger

Query Match      100.0%; Score 30; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  HORVH 5
      |||||
Db      47  HORVH 51

RESULT 4
I37963
zinc finger protein kox23 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 05-Nov-1999
C:Accession: I37963; S10412
R:Thiesen, H.J.
New Biol. 2, 363-374, 1990
A:Title: Multiple genes encoding zinc finger domains are expressed in human T cells.
A:Reference number: I37949; MUID:91145339; PMID:2288909
A:Accession: I37963
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-56 <RES>
A:Cross-references: EMBL:X52354; NID:g34156; PIDN:CAA36580.1; PID:q930086
C:Keywords: DNA binding; zinc finger
F:3-51/Region: zinc finger

Query Match      100.0%; Score 30; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  HORVH 5
      |||||
Db      47  HORVH 51

RESULT 5
I37966
zinc finger protein kox26 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 05-Nov-1999
C:Accession: I37966; S10415
R:Thiesen, H.J.
New Biol. 2, 363-374, 1990
A:Title: Multiple genes encoding zinc finger domains are expressed in human T cells.
A:Reference number: I37949; MUID:91145339; PMID:2288909
A:Accession: I37966
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-56 <RES>
A:Cross-references: EMBL:X52357; NID:g34159; PIDN:CAA36583.1; PID:q930089
C:Genetics:
A:Gene: GDB:2NF29
A:Cross-references: GDB:125336; OMIM:194535
A:Map position: 17pter-17p12
C:Keywords: DNA binding; zinc finger

Query Match      100.0%; Score 30; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  HORVH 5
      |||||
Db      47  HORVH 51

RESULT 6
S47065
finger protein HZF7, Krueppel-related - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C:Accession: S47065; I37573
R:Abbrink, M.; Aveskogh, M.; Hellman, L.
Submitted to the EMBL Data Library, June 1994
A:Description: Isolation of cDNA clones for 42 different Krueppel-related zinc finger
A:Reference number: I37566; MUID:95169271; PMID:7865130
A:Accession: S47065
A:Molecule type: mRNA
A:Residues: 1-66 <ABR>
R:Abbrink, M.; Aveskogh, M.; Hellman, L.
DNA Cell Biol. 14, 125-136, 1995
A:Title: Isolation of cDNA clones for 42 different Krueppel-related zinc finger prote
A:Reference number: I37573
A:Accession: I37573
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-66 <RES>
A:Cross-references: EMBL:X78930; NID:g498732; PIDN:CAA55530.1; PID:g836629
C:Genetics:
A:Gene: HZF7
C:Keywords: DNA binding; zinc finger

Query Match      100.0%; Score 30; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  HORVH 5
      |||||
Db      7  HORVH 11

RESULT 7
A42825
Krueppel-type zinc finger protein ZNF72 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C:Accession: A42825
R:Abry, M.; Marineau, C.; Zhang, F.F.; Zahed, L.; Figlewicz, D.; Delattre, O.; Thom
Genomics 13, 641-648, 1992
A:Title: Cloning of six new genes with zinc finger motifs mapping to short and long
A:Reference number: A42825; MUID:92347859; PMID:1639391
A:Accession: A42825
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-85 <AUB>
A>Note: sequence extracted from NCBI backbone (NCBIP:109770)
C:Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match      100.0%; Score 30; DB 2; Length 85;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 HQRVH 5
 Db 26 HQRVH 30

RESULT 8
 I38941
 zinc finger protein ZNF154 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 05-Nov-1999
 C:Accession: I38941
 R:Tommerup, N.; Vissing, H.
 Genomics 27, 259-264, 1995
 A:Title: Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identified by differential display
 A:Reference number: A57785; MUID:96044430; PMID:7557990
 A:Accession: I38941
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-87 <RES>
 A:Cross-references: EMBL:U20648; NID:9676874; PIDN:AAC50257.1; PID:9676875
 C:Genetics:
 A:Gene: GDB:ZNF154
 A:Cross-references: GDB:304989
 A:Map position: 19q13.4-19q13.3

Query Match 100.0%; Score 30; DB 2; Length 87;
 Best Local Similarity 100.0%; Pred. No. 6.3;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
 Db 25 HQRVH 29

RESULT 9
 D43284
 finger protein ZNF50 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1998
 C:Accession: D43284
 R:Lichter, P.; Bray, P.; Ried, T.; Dawid, I.B.; Ward, D.C.
 Genomics 13, 999-1007, 1992
 A:Title: Clustering of C2-H2 zinc finger motif sequences within telomeric and fragile sites
 A:Reference number: A43284; MUID:92372070; PMID:1505991
 A:Accession: D43284
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-88 <LIC>
 A:Cross-references: GB:M88360; NID:9340457; PID:9340458
 A:Note: sequence extracted from NCBI backbone (NCBIN:111638, NCBIPI:111639)
 C:Keywords: zinc finger

Query Match 100.0%; Score 30; DB 2; Length 88;
 Best Local Similarity 100.0%; Pred. No. 6.3;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
 Db 1 HQRVH 5

RESULT 10
 B48827
 zinc finger protein (clone CTfin92) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
 C:Accession: B48827
 R:Noe, T.; Fujiwara, Y.; Sezaki, M.; Fujimoto, H.; Higashinakagawa, T.
 Dev. Biol. 153, 356-367, 1992
 A:Title: Expression of a mouse zinc finger protein gene in both spermatocytes and oocytes
 A:Reference number: A48827; MUID:93012481; PMID:1397691
 A:Accession: B48827

A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-105 <NOC>
 A:Cross-references: GB:D10631; NID:g220644; PIDN:BAAO1481.1; PID:d1001956; PID:g220
 A:Experimental source: spermatogenic cells
 A:Note: sequence extracted from NCBI backbone (NCBIPI:114772)
 C:Keywords: DNA binding; transcription regulation

Query Match 100.0%; Score 30; DB 2; Length 105;
 Best Local Similarity 100.0%; Pred. No. 7.6;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
 Db 17 HQRVH 21

RESULT 11
 D31201
 GLI-related finger protein HRR2 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 02-Aug-1996
 C:Accession: D31201
 R:Ruppert, J.M.; Kinzler, K.W.; Wong, A.J.; Bigner, S.H.; Kao, F.T.; Law, M.L.; Seu
 Mol. Cell. Biol. 8, 3104-3113, 1988
 A:Title: The GLI-Kruppel family of human genes.
 A:Reference number: A93103; MUID:89096896; PMID:2850480
 A:Accession: D31201
 A:Molecule type: DNA
 A:Residues: 1-106 <RUP>
 C:Keywords: DNA binding; tandem repeat; zinc finger

Query Match 100.0%; Score 30; DB 2; Length 106;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
 Db 77 HQRVH 81

RESULT 12
 T34560
 hypothetical protein DKFP434J0650.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34560
 R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, October 1999
 A:Reference number: Z11540
 A:Accession: T34560
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-106 <POD>
 A:Cross-references: EMBL:AL122078
 A:Experimental source: adult testis; clone DKFP434J0650
 C:Genetics:
 A:Note: DKFP434J0650.1

Query Match 100.0%; Score 30; DB 2; Length 106;
 Best Local Similarity 100.0%; Pred. No. 7.7;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
 Db 44 HQRVH 48

RESULT 13
 Q085D9
 HHLF3 protein - human cytomegalovirus (strain AD169)
 N:Alternate names: hypothetical protein US33
 C:Species: human cytomegalovirus, human herpesvirus 5

C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 28-Jul-2000
C:Accession: I27216; S09948
R:Weston, K.; Barrell, B.G.
J. Mol. Biol. 192, 177-208, 1986
A:Title: Sequence of the short unique region, short repeats, and part of the long repeat
A:Reference number: A92935; MUID:87169717; PMID:3031311
A:Accession: I27216
A:Molecule type: DNA
A:Residues: 1-137 <RES>
A:Cross-references: EMBL:X04650; NID:g59801; PIDN:CAB37119.1; PID:g4456201
A:Experimental source: Strain AD169
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039; PMID:2161319
A:Accession: S09948
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-137 <CHE>
A:Cross-references: EMBL:X17403; NID:g59591; PIDN:CAA35266.1; PID:g1780965
A:Experimental source: strain AD169
A:Note: this sequence was submitted to the EMBL Data Library, December 1989
A:Note: this reading frame extends between two stop codons and does not begin with a sta
C:Genetics:
A:Gene: HHLF3
C:Superfamily: cytomegalovirus HHLF3 protein

Query Match 100.0%; Score 30; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
Db 72 HORVH 76

RESULT 14
I38620
zinc finger protein ZNF155 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 05-Nov-1999
C:Accession: I38620
R:Tommerup, N.; Vissing, H.
Genomics 27, 259-264, 1995
A:Title: Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identifi
A:Reference number: A57785; MUID:96044430; PMID:7557990
A:Accession: I38620
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-139 <RES>
A:Cross-references: EMBL:U09852; NID:g495575; PIDN:AAC50268.1; PID:g495576
C:Genetics:
A:Gene: GDB:ZNF155
A:Cross-references: GDB:304990
A:Map position: 19q13.2-19q13.32

Query Match 100.0%; Score 30; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
Db 65 HORVH 69

RESULT 15
A39240
finger protein mfg1 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 05-Nov-1999
C:Accession: A39240
R:Passananti, C.; Felsani, A.; Caruso, M.; Amati, P.

Proc. Natl. Acad. Sci. U.S.A. 86, 9417-9421, 1989
A:Title: Mouse genes coding for "zinc-finger"-containing proteins: characterization
A:Reference number: A39240; MUID:90083278; PMID:2512579
A:Accession: A39240
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <PAS>
A:Cross-references: GB:M28513; NID:g199136; PIDN:AAA39531.1; PID:g554203
C:Keywords: DNA binding; zinc finger

Query Match 100.0%; Score 30; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
Db 54 HORVH 58

Search completed: October 1, 2003, 10:47:52
Job time : 11.9722 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:06:46 ; Search time 5.83333 Seconds
(without alignments)
40.309 Million cell updates/sec

Title: US-10-057-890A-7

Perfect score: 30
Sequence: 1 HQRVH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	30	100.0	56	1 ZK23_HUMAN	P17034 homo sapien
2	30	100.0	56	1 ZN17_HUMAN	P17021 homo sapien
3	30	100.0	56	1 ZN26_HUMAN	P17031 homo sapien
4	30	100.0	56	1 ZN29_HUMAN	P17037 homo sapien
5	30	100.0	87	1 Z154_HUMAN	Q13106 homo sapien
6	30	100.0	105	1 ZF41_MOUSE	Q02526 mus musculu
7	30	100.0	106	1 HXR2_HUMAN	P10073 homo sapien
8	30	100.0	137	1 US33_HCMVA	P09697 human cytom
9	30	100.0	147	1 PRY_HUMAN	O14603 homo sapien
10	30	100.0	169	1 ZF58_MOUSE	P16372 mus musculu
11	30	100.0	172	1 ZG5A_XENLA	P18726 xenopus lae
12	30	100.0	196	1 ZG3_XENLA	P18718 xenopus lae
13	30	100.0	201	1 Z239_MOUSE	P24399 mus musculu
14	30	100.0	209	1 Z235_HUMAN	O14590 homo sapien
15	30	100.0	223	1 GLI4_HUMAN	P10075 homo sapien
16	30	100.0	223	1 ZG52_XENLA	P18727 xenopus lae
17	30	100.0	247	1 Z020_XENLA	P18744 xenopus lae
18	30	100.0	301	1 Z11A_HUMAN	P17013 homo sapien
19	30	100.0	348	1 Z134_HUMAN	P52741 homo sapien
20	30	100.0	367	1 Z211_HUMAN	Q13398 homo sapien
21	30	100.0	368	1 ZN24_HUMAN	P17028 homo sapien
22	30	100.0	368	1 ZN24_MOUSE	Q91vn1 mus musculu
23	30	100.0	379	1 Z11B_HUMAN	O06732 homo sapien
24	30	100.0	393	1 ZF46_MOUSE	Q03309 mus musculu
25	30	100.0	417	1 Z232_HUMAN	Q9uny5 homo sapien
26	30	100.0	420	1 Z229_HUMAN	Q9ujw7 homo sapien
27	30	100.0	445	1 ZG5_XENLA	P18725 xenopus lae
28	30	100.0	451	1 Z222_HUMAN	Q9uk12 homo sapien
29	30	100.0	458	1 Z239_HUMAN	Q16600 homo sapien
30	30	100.0	458	1 Z275_HUMAN	Q9nsd4 homo sapien
31	30	100.0	470	1 Z436_HUMAN	Q9c0f3 homo sapien
32	30	100.0	473	1 ZF38_HUMAN	Q9y5a6 homo sapien
33	30	100.0	474	1 Z230_HUMAN	Q9ue0d homo sapien

34	30	100.0	474	1 Z256_HUMAN	O9y2p7 homo sapien
35	30	100.0	482	1 Z223_HUMAN	Q9uk11 homo sapien
36	30	100.0	488	1 ZF92_MOUSE	Q62396 mus musculu
37	30	100.0	489	1 ZN71_HUMAN	O9nqz8 homo sapien
38	30	100.0	498	1 Z271_HUMAN	O14591 homo sapien
39	30	100.0	510	1 ZF64_HUMAN	P15622 homo sapien
40	30	100.0	518	1 SSG1_AVEA	P16615 avena sativ
41	30	100.0	518	1 SSG2_AVEA	P14812 avena sativ
42	30	100.0	521	1 Z286_HUMAN	Q9hbt8 homo sapien
43	30	100.0	534	1 Z397_HUMAN	O8nf99 homo sapien
44	30	100.0	536	1 Y628_HUMAN	O75123 homo sapien
45	30	100.0	538	1 Z155_HUMAN	Q12901 homo sapien

ALIGNMENTS

RESULT 1
ZK23_HUMAN
ID ZK23_HUMAN STANDARD; PRT; 56 AA.
AC P17034;
DI 01-AUG-1990 (Rel. 15, Created)
DI 01-AUG-1990 (Rel. 15, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein KOX23 (Fragment).
GN KOX23.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE=91145339; PubMed=2288909;
RA Thiesen H.-O.;
RT *Multiple genes encoding zinc finger domains are expressed in human T
RT cells.*
RL New Biol. 2:363-374(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X52354; CAA36580.1; -
CC PIR; I37963; I37963.
CC InterPro: IPR007087; znf_C2H2.
CC Pfam; PF000096; zf_C2H2; 2.
CC ProDom; PD000003; znf_C2H2; 2.
CC SMART; SM00355; znf_C2H2; 2.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
CC Zinc-finger; Metal-binding; DNA-binding.
CC NON_TER 1 1
CC ZN_FING 1 23 C2H2-TYPE.
CC ZN_FING 29 51 C2H2-TYPE.
CC NON_TER 56 56
CC SEQUENCE 56 AA; 6479 MW; B6BC3DEEB62D48CF CRC64;
Query Match 100.0%; Score 30; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 2,1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HQRVH 5
Db 47 HQRVH 51

RESULT 2
ZN17_HUMAN

```

ID 2N17_HUMAN STANDARD; PRT; 56 AA.
AC P17021;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 17 (Zinc finger protein KOX10) (HPF3) (Fragment).
GN ZNF17 OR KOX10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE-91145339; PubMed-2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
cells."
RL New Biol. 2:363-374(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION FACTOR.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC
DR EMBL; X52341; CAA36567.1; -
DR PIR; I37949; S10398.
DR Genew; HGNC:12958; ZNF17.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
Nuclear protein; Repeat.
FT NON_TER 1
FT ZN_FING 1 23 C2H2-TYPE.
FT ZN_FING 29 51 C2H2-TYPE.
FT NON_TER 56 56
FT SEQUENCE 56 AA; 6704 MW; CD18EF52E08F000C CRC64;
SQ
Query Match 100.0%; Score 30; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HQRVH 5
DB 19 HQRVH 23
RESULT 3
2N26_HUMAN STANDARD; PRT; 56 AA.
AC P17031;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 26 (Zinc finger protein KOX20) (Fragment).
GN ZNF26 OR KOX20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE-91145339; PubMed-2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
cells."
RL New Biol. 2:363-374(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC
DR EMBL; X52341; CAA36567.1; -
DR PIR; I37949; S10398.
DR Genew; HGNC:12958; ZNF17.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
Nuclear protein; Repeat.
FT NON_TER 1
FT ZN_FING 1 23 C2H2-TYPE.
FT ZN_FING 29 51 C2H2-TYPE.
FT NON_TER 56 56
FT SEQUENCE 56 AA; 6704 MW; CD18EF52E08F000C CRC64;
SQ
Query Match 100.0%; Score 30; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HQRVH 5
DB 19 HQRVH 23

```

```

RX MEDLINE-91145339; PubMed-2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
cells."
RL New Biol. 2:363-374(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC
DR EMBL; X52351; CAA36577.1; -
DR PIR; I37960; I37960.
DR HSSP; P08045; 12NF.
DR Genew; HGNC:13053; ZNF26.
DR MIM; 194537; -
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW Transcription regulation; Zinc-finger; DNA-binding; Metal-binding;
Nuclear protein; Repeat.
FT NON_TER 1
FT ZN_FING 1 23 C2H2-TYPE.
FT ZN_FING 29 51 C2H2-TYPE.
FT NON_TER 56 56
FT SEQUENCE 56 AA; 6408 MW; 31ACBC52B8BF2855F CRC64;
SQ
Query Match 100.0%; Score 30; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HQRVH 5
DB 47 HQRVH 51
RESULT 4
2N29_HUMAN STANDARD; PRT; 56 AA.
AC P17037;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 29 (Zinc finger protein KOX26) (Fragment).
GN ZNF29 OR KOX26.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE-91145339; PubMed-2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
cells."
RL New Biol. 2:363-374(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC
DR EMBL; X52351; CAA36577.1; -
DR PIR; I37960; I37960.
DR HSSP; P08045; 12NF.
DR Genew; HGNC:13053; ZNF26.
DR MIM; 194537; -
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW Transcription regulation; Zinc-finger; DNA-binding; Metal-binding;
Nuclear protein; Repeat.
FT NON_TER 1
FT ZN_FING 1 23 C2H2-TYPE.
FT ZN_FING 29 51 C2H2-TYPE.
FT NON_TER 56 56
FT SEQUENCE 56 AA; 6408 MW; 31ACBC52B8BF2855F CRC64;
SQ
Query Match 100.0%; Score 30; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HQRVH 5
DB 47 HQRVH 51

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X52357; CAA36583.1; -
DR PIR; I37966; I37966.
DR Genew; HGNC:13080; ZNF29.
DR MIM; 194535; -
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW Transcription regulation; Zinc-finger; DNA-binding; Metal-binding;
KW Nuclear protein; Repeat.
FT NON_TER 1 1
FT ZN_FING 1 23 C2H2-TYPE.
FT ZN_FING 29 51 C2H2-TYPE.
FT NON_TER 56 56
SQ SEQUENCE 56 AA: 6368 MW: D812CB5BF419884 CRC64;

Query Match 100.0%; Score 30; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 47 HORVH 51

RESULT 5
Z154 HUMAN
ID Z154_HUMAN STANDARD; PRT; 87 AA.
AC Q13106;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 154 (Fragment).
GN ZNF154.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Insulinoma;
RX MEDLINE=96044430; PubMed=7557990;
RA Tommerup N., Vissing H.;
RT *Isolation and fine mapping of 16 novel human zinc finger-encoding
RT cDNAs identify putative candidate genes for developmental and
RT malignant disorders.*;
RL Genomics 27:259-264(1995).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U20648; AAC50257.1; -
DR PIR; I38941; I38941.
DR HSP; P08047; ISP2.
DR Genew; HGNC:12939; ZNF154.
DR MIM; 604085; -

DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0003355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINCFINGER.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; Zinc-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT NON_TER 1 1
FT ZN_FING 7 29 C2H2-TYPE.
FT ZN_FING 35 57 C2H2-TYPE.
FT ZN_FING 63 85 C2H2-TYPE.
SQ SEQUENCE 87 AA: 10070 MW: 42B9C4E1AC9A92B3 CRC64;

Query Match 100.0%; Score 30; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 25 HORVH 29

RESULT 6
ZF41_MOUSE
ID ZF41_MOUSE STANDARD; PRT; 105 AA.
AC Q02526;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 41 (Zfp-41) (CtFIN92) (Fragment).
GN ZFP41 OR ZFP-41.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BTBR; TISSUE=Spermatocyte;
RX MEDLINE=93012481; PubMed=1397691;
RA Noce T., Fujiwara Y., Sezaki M., Fujimoto H., Higashinakagawa T.;
RT *Expression of a mouse zinc finger protein gene in both spermatocytes
RT and oocytes during meiosis.*;
RL Dev. Biol. 153:356-367(1992).
CC -!- FUNCTION: A PUTATIVE DNA-BINDING REGULATORY PROTEIN ASSOCIATED
CC WITH MEIOSIS IN SPERMATOGENESIS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN THE SPERMATOCYTES AND
CC SPERMATIDS OF TESTES. IT IS ALSO EXPRESSED IN THE FETUS AND
CC EMBRYONIC STEM CELLS AT LOWER LEVELS.
CC -!- DEVELOPMENTAL STAGE: DETECTED IN THE NEWBORN TESTIS AND PEAKS
CC AT 3 WEEKS DURING THE FIRST CYCLE OF SPERMATOGENESIS. EXPRESSED
CC IN THE FETUS AND EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D10631; BAA01481.1; -
DR PIR; B48827; B48827.
DR HSP; P08046; IALI.
DR MGD; MGI:99186; Zfp41.
DR InterPro; IPR007087; Znf_C2H2.

```

DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF00096; zf-C2H2; 4.
DR PRINTS; PR00048; ZINC_FINGER.
DR SMART; SM00355; Znf_C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 4.
DR Zinc-finger; Metal-binding; Developmental protein; Spermatogenesis;
KW Nuclear protein; Transcription regulation; DNA-binding; Repeat.
FT NON_TER 1
FT ZN_FING 1 20 C2H2-TYPE.
FT ZN_FING 29 49 C2H2-TYPE.
FT ZN_FING 57 76 C2H2-TYPE.
FT ZN_FING 85 105 C2H2-TYPE.
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 12083 MW; 2A29456BB9A7B1F CRC64;

Query Match 100.0%; Score 30; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 17 HORVH 21

RESULT 7
HKR2_HUMAN STANDARD; PRT; 106 AA.
AC P10073;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Krueppel-related zinc finger protein 2 (HKR2 protein) (Fragment).
GN HKR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89096896; PubMed=2850480;
RA Ruppert J.M., Kinzler K.W., Wong A.J., Bigner S.H., Kao F.T.,
RA Law M.L., Seunavaz H.N., O'Brien S.J., Vogelstein B.;
RT "The Gli-Kruppel family of human genes.";
RL Mol. Cell. Biol. 8:3104-3113(1988).
CC !- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.
CC !- SUBCELLULAR LOCATION: Nuclear (Potential).
CC !- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20676; -; NOT_ANNOTATED_CDS.
CC PIR; D31201; D31201.
CC RSP; P08045; 12NF.
CC Genew; HGNC:4929; HKR2.
CC MIM; 165260; -.
CC GO; GO:0005634; C:nucleus; NAS.
CC InterPro; IPR007087; Znf_C2H2.
CC InterPro; IPR007086; Znf_C2H2_sub.
CC Pfam; PF00096; zf-C2H2; 3.
CC PRINTS; PR00048; ZINC_FINGER.
CC SMART; SM00355; Znf_C2H2; 3.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 4.
CC Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.

```

```

FT NON_TER 1 1
FT ZN_FING 7 25 C2H2-TYPE.
FT ZN_FING 31 54 C2H2-TYPE.
FT ZN_FING 59 82 C2H2-TYPE.
FT ZN_FING 87 >106 C2H2-TYPE.
FT NON_TER 106 106
SQ SEQUENCE 106 AA; 12068 MW; 95A57A8EA865B403 CRC64;

Query Match 100.0%; Score 30; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 77 HORVH 81

RESULT 8
US33_HCMVA STANDARD; PRT; 137 AA.
ID US33_HCMVA
AC P09697;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical protein HHLF3.
GN US33.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87169717; PubMed=3031311;
RA Weston K., Barrell B.G.;
RT "Sequence of the short unique region, short repeats, and part of the
RT long repeats of human cytomegalovirus.";
RL J. Mol. Biol. 192:177-208(1986).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17403; CAA35266.1; -
CC EMBL; X04650; CAB37119.1; -
CC PIR; I27216; Q0BED9.
CC KW Hypothetical protein.
CC SEQUENCE 137 AA; 15774 MW; 17DE2F61C2A859B5 CRC64;

Query Match 100.0%; Score 30; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 72 HORVH 76

RESULT 9
PRY_HUMAN STANDARD; PRT; 147 AA.
ID PRY_HUMAN

```

AC O14603;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE PTPN13-like protein, Y-linked (Testis-specific PTP-BL related Y
 GN protein).
 GN PRT OR PTPN13LY.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Testis;
 RX MEDLINE=98022381; PubMed=9381176;
 RA Lahn B.T., Page D.C.;
 RT "Functional coherence of the human Y chromosome.";
 RL Science 278:675-680(1997).
 RN [2]
 RP REVISION.
 RA Lahn B.T., Page D.C.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Kuroda-Kawaguchi T., Skaletsky H., Minx P.J., Brown L.G., Rozen S.,
 RA Wilson R.K., Waterston R.H., Page D.C.;
 RT "The DNA sequence of the human Y chromosome.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=O14603-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O14603-2; Sequence=VSP_004068, VSP_004069;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Testis.
 CC -1- CAUTION: PRT HAS MULTIPLE LOCATIONS ON CHROMOSOME Y.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF000988; AAC51835.2; -;
 DR EMBL; AF517635; AN06674.1; -;
 DR Genew; HGNC:14024; PRT;
 DR MIM; 400019; -;
 KW Alternative splicing.
 FT VARSPLIC 63 71 YKRVGCTSL -> DGVSPCLPC (in isoform 2).
 FT VARSPLIC 63 71 /FTid=VSP_004068.
 FT VARSPLIC 72 147 Missing (in isoform 2).
 FT VARSPLIC 72 147 /FTid=VSP_004069.
 SQ SEQUENCE 147 AA: 16512 MW: 87C861FC2BEDAE9B CRC64;
 Query Match 100.0%; Score 30; DB 1; Length 147;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HQRVH 5
 DB 93 HQRVH 97
 RESULT 10
 ZF58_MOUSE
 ID ZF58_MOUSE STANDARD; PRT; 169 AA.
 AC P16372;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Zinc finger protein 58 (zfp-58) (Zinc finger protein Mfg-1)
 DE (Fragment).
 GN ZFP58 OR MFG-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Skeletal muscle;
 RX MEDLINE=90083278; PubMed=2512579;
 RA Passananti C., Felsani A., Caruso M., Amati P.;
 RT "Mouse genes coding for 'zinc-finger'-containing proteins:
 RT characterization and expression in differentiated cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:9417-9421(1989).
 CC -1- FUNCTION: MAY HAVE A ROLE DURING DIFFERENTIATION PROCESSES.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: LIVER; TESTIS; AND AT CONSIDERABLY LOWER
 CC LEVELS IN BRAIN; SPLEEN; HEART.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS POSITIVELY REGULATED UPON
 CC DIFFERENTIATION, AND IS NOT RELATED TO THE CELL CYCLE.
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE 2INC-
 CC FINGER PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M28513; AAA39531.1; -;
 DR PIR; A39240; A39240.
 DR HSP; P25490; IUBD.
 DR MGD; MGI:99205; Zfp58.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF00096; zf-C2H2; 6.
 DR PRINTS; PR00048; ZNCFINGER.
 DR SMART; SM00355; Znf_C2H2; 6.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.
 FT NON_TER 1 1
 FT ZN_FING 8 30 C2H2-TYPE.
 FT ZN_FING 36 58 C2H2-TYPE.
 FT ZN_FING 64 86 C2H2-TYPE.
 FT ZN_FING 92 114 C2H2-TYPE.
 FT ZN_FING 120 142 C2H2-TYPE.
 FT ZN_FING 148 >169 C2H2-TYPE.
 FT NON_TER 169 169
 SQ SEQUENCE 169 AA: 19585 MW: 7F5256B1406CDA3 CRC64;
 Query Match 100.0%; Score 30; DB 1; Length 169;
 Best Local Similarity 100.0%; Pred. No. 6.7;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HQRVH 5
 DB 54 HQRVH 58
 RESULT 11
 ZG5A_XENLA
 ID ZG5A_XENLA STANDARD; PRT; 172 AA.
 AC P18726;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Gastrula zinc finger protein XLCGF51.1A (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 RN NCBI_TaxID=8355;
 RN [1]
 RX MEDLINE=90040698; PubMed=2509712;
 RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
 RA Poetling A., Knoechel W.,
 RT "Second-order repeats in Xenopus laevis finger proteins.";
 RL J. Mol. Biol. 208:639-659(1989).
 RL PIR; S06575; S06575.
 DR HSSP; P03001; 1TF3.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; Znf_C2H2; 6.
 DR SMART; SM00355; Znf_C2H2; 6.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
 FT NON_TER 1
 FT ZN_FING 6 28 C2H2-TYPE.
 FT ZN_FING 34 56 C2H2-TYPE.
 FT ZN_FING 62 84 C2H2-TYPE.
 FT ZN_FING 90 112 C2H2-TYPE.
 FT ZN_FING 122 144 C2H2-TYPE.
 FT ZN_FING 150 172 C2H2-TYPE.
 FT NON_TER 172 172
 SQ SEQUENCE 172 AA; 19609 MW; AF8290211AB5E0D0 CRC64;

 Query Match 100.0%; Score 30; DB 1; Length 172;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 HQRVH 5
 Db 168 HQRVH 172

 RESULT 12
 ZG3_XENLA STANDARD; PRT; 196 AA.
 AC P18718;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Gastrula zinc finger protein XICGF3.1 (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 RN NCBI_TaxID=8355;
 RN [1]
 RX MEDLINE=90040698; PubMed=2509712;
 RA Nietfeld W., El-Baradi T., Mentzel H., Pieler T., Koester M.,
 RA Poetling A., Knoechel W.,
 RT "Second-order repeats in Xenopus laevis finger proteins.";
 RL J. Mol. Biol. 208:639-659(1989).
 RL PIR; S06575; S06575.
 DR HSSP; P08047; ISP2.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; Znf_C2H2; 7.
 DR SMART; SM00355; Znf_C2H2; 7.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 7.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
 FT NON_TER 1
 FT ZN_FING 6 28 C2H2-TYPE.
 FT ZN_FING 34 56 C2H2-TYPE.
 FT ZN_FING 62 84 C2H2-TYPE.
 FT ZN_FING 90 112 C2H2-TYPE.
 FT ZN_FING 118 140 C2H2-TYPE.
 FT ZN_FING 146 168 C2H2-TYPE.
 FT ZN_FING 174 196 C2H2-TYPE.

FT NON_TER 196 196
 SQ SEQUENCE 196 AA; 22378 MW; B5735306E722FFDC CRC64;

 Query Match 100.0%; Score 30; DB 1; Length 196;
 Best Local Similarity 100.0%; Pred. No. 7.8;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 HQRVH 5
 Db 136 HQRVH 140

 RESULT 13
 Z239_MOUSE STANDARD; PRT; 201 AA.
 ID 2239_MOUSE
 AC P24399;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Zinc finger protein 239 (Zfp-239) (Zinc finger protein MOK-2).
 DE ZNF239 OR ZFP239 OR MOK2 OR MOK-2.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=90097859; PubMed=2104662;
 RA Ernault-Lange M., Kress M., Hamer D.;
 RT "A gene that encodes a protein consisting solely of zinc finger
 domains is preferentially expressed in transformed mouse cells.";
 RL Mol. Cell. Biol. 10:418-421(1990).
 CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN TRANSFORMED
 CC MOUSE CELLS.
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -!- SIMILARITY: Contains 7 C2H2-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M32057; AAA39729.1; -;
 DR PIR; I57505; I57505.
 DR HSSP; P08047; ISP2.
 DR TRANSFAC; T00510; -;
 DR MGD; MGI:1306812; Zfp239.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR007086; Znf_C2H2_sub.
 DR Pfam; PF00096; Znf_C2H2; 7.
 DR PRINTS; PR00046; ZINC_FINGER.
 DR ProDom; PD000003; Znf_C2H2; 6.
 DR SMART; SM00355; Znf_C2H2; 7.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 7.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat.
 FT NON_TER 6 28 C2H2-TYPE 1.
 FT ZN_FING 34 56 C2H2-TYPE 2.
 FT ZN_FING 62 84 C2H2-TYPE 3.
 FT ZN_FING 90 112 C2H2-TYPE 4.
 FT ZN_FING 118 140 C2H2-TYPE 5.
 FT ZN_FING 146 168 C2H2-TYPE 6.
 FT ZN_FING 174 196 C2H2-TYPE 7.
 SQ SEQUENCE 201 AA; 22832 MW; B7A18C0AADAB1384 CRC64;

 Query Match 100.0%; Score 30; DB 1; Length 201;

```
Best Local Similarity 100.0%; Pred. No. 8;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 80 HORVH 84

RESULT 14
ID 2335_HUMAN STANDARD; PRT; 209 AA.
AC Q14590;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 935 (Zinc finger protein 93 homolog) (Zfp-93)
GN (Zinc finger protein HZF6) (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=95169271; PubMed=7865130;
RA Adirink M., Aveskogh M., Hellman L.;
RT "Isolation of cDNA clones for 42 different Kruppel-related zinc finger
RL DNA Cell Biol. 14:125-136(1995).
CC -!- FUNCTION: May function as a transcription factor.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X78929; CAA55529.1; -
DR PIR: S47068; S47068.
DR HSSP: P25490; 12NM.
DR Genew; HGNC:12866; ZNF235.
DR MIM; 604749; -
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 7.
DR ProDom; PD000003; Znf_C2H2; 6.
DR SMART; SM00355; Znf_C2H2; 7.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 7.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT NON_TER 1 1
FT ZN_FING 4 26 C2H2-TYPE.
FT ZN_FING 32 54 C2H2-TYPE.
FT ZN_FING 60 82 C2H2-TYPE.
FT ZN_FING 88 110 C2H2-TYPE.
FT ZN_FING 116 138 C2H2-TYPE.
FT ZN_FING 144 166 C2H2-TYPE.
FT ZN_FING 172 194 C2H2-TYPE.
FT ZN_FING 200 >209 C2H2-TYPE.
FT NON_TER 209 209
SQ SEQUENCE 209 AA; 23915 MW; 9C12DE06BC7FFC8E CRC64;

Query Match 100.0%; Score 30; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 80 HORVH 84
```

```
Db 22 HORVH 26

RESULT 15
ID GLI4_HUMAN STANDARD; PRT; 223 AA.
AC P10075;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE GLI4 protein (Kruppel-related zinc finger protein 4) (HKR4 protein)
DE (Fragments).
GN GLI4 OR HKR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=89096896; PubMed=2850480;
RA Ruppert J.M., Kinzler K.W., Wong A.J., Bigner S.H., Kao F.T.,
RA Law M.L., Seunanez H.N., O'Brien S.J., Vogelstein B.;
RT "The GLI-Kruppel family of human genes.";
RL Mol. Cell. Biol. 8:3104-3113(1988).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20678; AAA35990.1; -
DR EMBL; M20679; -; NOT_ANNOTATED_CDS.
DR PIR: F31201; F31201.
DR HSSP; P08046; 1AIG.
DR Genew; HGNC:4320; GLI4.
DR MIM; 165280; -
DR GO; GO:0005634; C:nucleus; NAS.
DR InterPro; IPR007087; Znf_C2H2.
DR SMART; SM00355; Znf_C2H2; 6.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.
KW Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat.
FT NON_TER 1 1
FT ZN_FING 40 62 C2H2-TYPE.
FT ZN_FING 68 90 C2H2-TYPE.
FT ZN_FING 96 118 C2H2-TYPE.
FT ZN_FING 124 >129 C2H2-TYPE.
FT NON_CONS 129 130
FT ZN_FING <130 136 C2H2-TYPE.
FT ZN_FING 142 164 C2H2-TYPE.
FT ZN_FING 170 192 C2H2-TYPE.
FT ZN_FING 198 220 C2H2-TYPE.
SQ SEQUENCE 223 AA; 25276 MW; 5DF24D508A04EA69 CRC64;

Query Match 100.0%; Score 30; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 216 HORVH 220

Search completed: October 1, 2003, 10:42:42
Job time : 6.83333 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:33:32 ; Search time 29.8611 Seconds
(without alignments)
43.209 Million cell updates/sec

Title: US-10-057-890A-7
Perfect score: 30
Sequence: 1 HQRVH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rhodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Query %		Length	DB	ID	Description
	Score	Match				
1	30	100.0	48	11	P97663	P97663 rattus norv
2	30	100.0	53	4	O14850	O14850 homo sapien
3	30	100.0	57	11	O88233	O88233 mus musculu
4	30	100.0	57	11	O9QW47	O9QW47 mus musculu
5	30	100.0	57	11	O9QW40	O9QW40 mus musculu
6	30	100.0	70	4	O14886	O14886 homo sapien
7	30	100.0	74	4	O14887	O14887 homo sapien
8	30	100.0	74	4	O14888	O14888 homo sapien
9	30	100.0	75	4	O14890	O14890 homo sapien
10	30	100.0	85	11	O88258	O88258 mus musculu
11	30	100.0	85	11	O88253	O88253 mus musculu
12	30	100.0	88	4	O15922	O15922 homo sapien
13	30	100.0	92	11	O8C288	O8C288 mus musculu
14	30	100.0	102	4	O14899	O14899 homo sapien
15	30	100.0	106	4	O9UFH1	O9UFH1 homo sapien
16	30	100.0	107	11	O60916	O60916 mus musculu

17 30 100.0 130 6 Q95286
18 30 100.0 131 16 Q8DA53
19 30 100.0 137 6 Q29240
20 30 100.0 145 4 Q8NB43
21 30 100.0 151 6 Q9XSR2
22 30 100.0 153 4 Q8TCQ7
23 30 100.0 153 16 Q8PJ7
24 30 100.0 153 16 Q8P892
25 30 100.0 167 4 Q8N932
26 30 100.0 175 11 Q9ES79
27 30 100.0 179 11 Q9JJR2
28 30 100.0 184 4 Q15914
29 30 100.0 198 4 Q8N8Y5
30 30 100.0 198 11 Q8BSV7
31 30 100.0 199 4 Q9NWL3
32 30 100.0 201 11 Q99J71
33 30 100.0 205 4 Q15918
34 30 100.0 205 4 Q9H707
35 30 100.0 206 5 Q9VEK2
36 30 100.0 207 4 Q8NDL8
37 30 100.0 208 4 Q8N4Z8
38 30 100.0 229 4 Q8NS06
39 30 100.0 230 4 Q9BTG0
40 30 100.0 230 4 Q96IT2
41 30 100.0 251 16 Q8PJ01
42 30 100.0 251 16 Q8P896
43 30 100.0 252 4 Q8NAS5
44 30 100.0 258 16 O54125
45 30 100.0 261 4 Q8N3J9

ALIGNMENTS

RESULT 1

P97663 ID P97663 PRELIMINARY; PRT; 48 AA.
AC P97663;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Zinc finger protein 7 (Fragment).
GN DZF7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96319713; PubMed=8697448;
RA Gebelein B., Mesa K., Urrutia R.;
RT "A novel profile of expressed sequence tags for zinc finger encoding genes from the poorly differentiated exocrine pancreatic cell line AR4IP."
RL Cancer Lett. 105:225-231(1996).
DR EMBL; U78133; AAB36807.1; -;
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 1.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW Metal-binding; zinc; zinc-finger.
FT NON_TER 1 1
FT NON_TER 48 48
SQ SEQUENCE 48 AA; 5388 MW; 3A5A7EA34459B06A CRC64;

Query Match 100.0%; Score 30; DB 11; Length 48;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HQRVH 5
|||||

```

Db          39 HORVH 43

RESULT 2
O14850      PRELIMINARY;      PRT;      53 AA.
AC O14850;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Zinc finger protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RU Wu G., Yu L., Sun X., Wu M., Fan Y., Jiang C., Zheng Q.,
RA Zhao S.;
RT *Isolation and Cloning of Novel C2-H2 Type Zinc Finger Protein Gene in
RL Human Liver Tissue"; to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF024698; AAB81088.1; -.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF000096; zf-C2H2; 1.
DR PRINTS; PR00048; ZINCFINGER.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 53
SQ SEQUENCE 53 AA: 6029 MR; 783F4B143C4AD012 CRC64;

Query Match 100.0%; Score 30; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 6 HORVH 10

RESULT 3
O88233      PRELIMINARY;      PRT;      57 AA.
AC O88233;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Zinc finger protein (Fragment).
GN MSZF16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL6;
RX MEDLINE=98296253; PubMed=9630514;
RA Agata Y., Matsuda E., Shimizu A.;
RT *Rapid and efficient cloning of cDNAs encoding krueppel-like zinc
RL finger proteins by degenerate PCR.";
DR EMBL; AB010330; BAA31386.1; -.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF000096; zf-C2H2; 2.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
FT NON_TER 1
FT NON_TER 57
SQ SEQUENCE 57 AA: 6789 MW; 88D5E406C437CA1E CRC64;

Query Match 100.0%; Score 30; DB 11; Length 57;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 50 HORVH 54

RESULT 5
O9QWN0      PRELIMINARY;      PRT;      57 AA.
AC O9QWN0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mszf28 (Fragment).
GN MSZF28.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
KW Metal-binding; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 57
SQ SEQUENCE 57 AA: 6468 MW; 01E2DD7DA7E99568 CRC64;

Query Match 100.0%; Score 30; DB 11; Length 57;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 50 HORVH 54

RESULT 4
O9QWN7      PRELIMINARY;      PRT;      57 AA.
AC O9QWN7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mszf92 (Fragment).
GN MSZF92.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL6;
RX MEDLINE=98296253; PubMed=9630514;
RA Agata Y., Matsuda E., Shimizu A.;
RT *Rapid and efficient cloning of cDNAs encoding Krueppel-like zinc
RL finger proteins by degenerate PCR.";
DR EMBL; AB010356; BAA31412.1; -.
DR HSSP; P08045; 1ZNF.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF000096; zf-C2H2; 2.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 2.
KW Metal-binding; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 57
SQ SEQUENCE 57 AA: 6789 MW; 88D5E406C437CA1E CRC64;

Query Match 100.0%; Score 30; DB 11; Length 57;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 50 HORVH 54

RESULT 5
O9QWN0      PRELIMINARY;      PRT;      57 AA.
AC O9QWN0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mszf28 (Fragment).
GN MSZF28.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-C57BL6;
 RX MEDLINE-98296253; PubMed-9630514;
 RA Agata Y., Matsuda E., Shimizu A.;
 RT "Rapid and efficient cloning of cDNAs encoding Krueppel-like zinc
 RT finger proteins by degenerate PCR.";
 RL Gene 213:55-64(1998).
 DR EMBL; AB010328; BAA31384.1; -.
 DR HSSP; P08047; 1SP2.
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam: PF00096; zf-C2H2; 2.
 DR ProDom: PD000003; Znf_C2H2; 2.
 DR SMART; SM00355; Znf_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
 KW Metal-binding; Zinc; Zinc-finger.
 FT NON_TER 1 57
 FT NON_TER 57 57
 SQ SEQUENCE 57 AA; 6611 MW; D61698C3A649B8F6 CRC64;

Query Match 100.0%; Score 30; DB 11; Length 57;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5

Db 22 HQRVH 26

RESULT 6

OL4886 ID OL4886 PRELIMINARY; PRT; 70 AA.
 AC OL4886;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Zinc finger protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RA Sun X., Yu L., Wu G., Liu S., Zhang M., Jiang Y., Shouyuan S.;
 RT "Isolation of novel human genes coding zinc finger protein from brain
 RT tissue - according to the conservativity of zinc finger motif.";
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 0:0-0(1997).
 DR EMBL; AF027141; AAB84020.1; -.
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 2.
 DR ProDom; PD000003; Znf_C2H2; 1.
 DR SMART; SM00355; Znf_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
 KW Metal-binding; Zinc; Zinc-finger.
 FT NON_TER 1 70
 FT NON_TER 70 70
 SQ SEQUENCE 70 AA; 8024 MW; BC3163DAD22E1FAB CRC64;

Query Match 100.0%; Score 30; DB 4; Length 70;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5

Db 25 HQRVH 29

RESULT 7

OL4887 ID OL4887 PRELIMINARY; PRT; 74 AA.
 AC OL4887;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Zinc finger protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-72 FROM N.A.
 RC TISSUE-Brain;
 RA Sun X., Yu L., Wu G., Liu S., Zhang M., Jiang Y., Shouyuan S.;
 RT "Isolation of novel human genes coding zinc finger protein from brain
 RT tissue - according to the conservativity of zinc finger motif.";
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 0:0-0(1997).
 DR EMBL; AF027142; AAB84021.1; -.
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 2.
 DR ProDom; PD000003; Znf_C2H2; 2.
 DR SMART; SM00355; Znf_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
 KW Metal-binding; Zinc; Zinc-finger.
 FT NON_TER 1 74
 FT NON_TER 74 74
 SQ SEQUENCE 74 AA; 8452 MW; FD75F15BD0635243 CRC64;
 Query Match 100.0%; Score 30; DB 4; Length 74;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5

Db 29 HQRVH 33

RESULT 8

OL4888 ID OL4888 PRELIMINARY; PRT; 74 AA.
 AC OL4888;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Zinc finger protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-72 FROM N.A.
 RC TISSUE-Brain;
 RA Sun X., Yu L., Wu G., Liu S., Zhang M., Jiang Y., Shouyuan S.;
 RT "Isolation of novel human genes coding zinc finger protein from brain
 RT tissue - according to the conservativity of zinc finger motif.";
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 0:0-0(1997).
 DR EMBL; AF027143; AAB84022.1; -.
 DR InterPro: IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 2.
 DR ProDom; PD000003; Znf_C2H2; 2.
 DR SMART; SM00355; Znf_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
 KW Metal-binding; Zinc; Zinc-finger.
 FT NON_TER 1 74
 FT NON_TER 74 74
 SQ SEQUENCE 74 AA; 8452 MW; FD75F15BD0635243 CRC64;

```

FT  NON_TER      74
SQ  SEQUENCE      74 AA;  8564 MW;  6F8711436CC847B6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 30; DB 4; Length 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 HORVH 5
DB  29 HORVH 33

RESULT 9
O14890
ID  O14890      PRELIMINARY;      PRT;      75 AA.
AC  O14890;
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Zinc finger protein (Fragment).
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE OF 1-72 FROM N.A.
RC  TISSUE=Brain;
RA  Sun X., Yu L., Wu G., Liu S., Hu P., Zhang M., Jiang Y., Shouyuan S.;
RT  "Isolation of novel human genes coding zinc finger protein from brain
RT  tissue - according to the conservativity of zinc finger motif.";
RL  Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 0:0-0(1997).
[2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RA  Zhang Q.;
RL  Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DE  EMBL: AF021745; AA84024.1; -.
DR  InterPro; IPR007087; Znf_C2H2.
DR  Pfam; PF00096; zf-C2H2; 2.
DR  ProDom; PD000003; Znf_C2H2; 2.
DR  SMART; SM00355; Znf_C2H2; 2.
DR  PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR  PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW  Metal-binding; Zinc; Zinc-finger.
FT  NON_TER      1
FI  NON_TER      75
SQ  SEQUENCE      75 AA;  8377 MW;  762646DE0AC4E752 CRC64;

Query Match
Best Local Similarity 100.0%; Score 30; DB 4; Length 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 HORVH 5
DB  30 HORVH 34

RESULT 10
O88258
ID  O88258      PRELIMINARY;      PRT;      85 AA.
AC  O88258;
DT  01-NOV-1998 (TrEMBLrel. 08, Created)
DT  01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Mszf51 (Fragment).
GN  MSZF51.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL6;

```

```

RX  MEDLINE=98296253; PubMed=9630514;
RA  Agata Y., Matsuda E., Shimizu A.;
RT  "Rapid and efficient cloning of cDNAs encoding Krueppel-like zinc
RT  finger proteins by degenerate PCR.";
RL  Gene 213:55-64(1998).
DR  EMBL; AB010363; BAA31419.1; -.
DR  InterPro; IPR007087; Znf_C2H2.
DR  Pfam; PF00096; zf-C2H2; 3.
DR  SMART; SM00355; Znf_C2H2; 3.
DR  PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR  PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW  Metal-binding; Zinc; Zinc-finger.
FT  NON_TER      1
FI  NON_TER      85
SQ  SEQUENCE      85 AA;  9681 MW;  CD65F331C63F5887 CRC64;

Query Match
Best Local Similarity 100.0%; Score 30; DB 11; Length 85;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 HORVH 5
DB  50 HORVH 54

RESULT 11
O88253
ID  O88253      PRELIMINARY;      PRT;      85 AA.
AC  O88253;
DT  01-NOV-1998 (TrEMBLrel. 08, Created)
DT  01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Mszf39 (Fragment).
DE  ZFP51 OR MSZF39.
GN  ZFP51 OR MSZF39.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL6;
RX  MEDLINE=98296253; PubMed=9630514;
RA  Agata Y., Matsuda E., Shimizu A.;
RT  "Rapid and efficient cloning of cDNAs encoding Krueppel-like zinc
RT  finger proteins by degenerate PCR.";
RL  Gene 213:55-64(1998).
DR  EMBL; AB010357; BAA31413.1; -.
DR  HSP; P08045; 12NF.
DR  MGD; MGI:99198; Zfp51.
DR  InterPro; IPR007087; Znf_C2H2.
DR  Pfam; PF00096; zf-C2H2; 3.
DR  ProDom; PD000003; Znf_C2H2; 3.
DR  SMART; SM00355; Znf_C2H2; 3.
DR  PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR  PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW  Metal-binding; Zinc; Zinc-finger.
FT  NON_TER      1
FI  NON_TER      85
SQ  SEQUENCE      85 AA;  10129 MW;  5CE07A6F82393A1A CRC64;

Query Match
Best Local Similarity 100.0%; Score 30; DB 11; Length 85;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 HORVH 5
DB  50 HORVH 54

RESULT 12
Q15922
ID  Q15922      PRELIMINARY;      PRT;      88 AA.
AC  Q15922;

```

DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE DNA-binding protein (Fragment).
 GN ZNF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=92052132; PubMed=1946370;
 RA Bray P., Lichter P., Thiesen H.J., Ward D.C., David I.B.;
 RT "Characterization and mapping of human genes encoding zinc finger
 RI proteins";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9563-9567(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=92372070; PubMed=1505991;
 RA Lichter P., Bray P., Ried T., David I.B., Ward D.C.;
 RT "Clustering of C2-H2 zinc finger motif sequences within telomeric and
 RI fragile site regions of human chromosomes";
 RL Genomics 13:999-1007(1992).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: M88360; AAA61318.1; -.
 DR HSSP: P08045; 1ZNF.
 DR InterPro: IPR007087; Znf_C2H2.
 DR ProDom: PD000003; Znf_C2H2_sub.
 DR InterPro: IPR007086; Znf_C2H2_sub.
 DR Pfam: PF00096; zf-C2H2; 3.
 DR PRINTS: PR00048; ZINCFINGER.
 DR ProDom: PD000003; Znf_C2H2; 1.
 DR SMART: SM00355; Znf_C2H2; 3.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 3.
 KW DNA-binding; Metal-binding; Nuclear protein; Zinc; Zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 88 AA; 10007 MW; 3C5A45BEA464760E CRC64;

 Query Match 100.0%; Score 30; DB 4; Length 88;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 HORVH 5
 Db |||||
 1 HORVH 5

 RESULT 13
 Q8C2B8 PRELIMINARY; PRT; 92 AA.
 ID Q8C2B8
 AC Q8C2B8;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Zinc finger protein 191 (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA The RIKEN Genome Exploration Research Group Phase 1 & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 DR EMBL: AK088914; BAC40649.1; -.
 FT NON_TER 1
 SQ SEQUENCE 92 AA; 10489 MW; 072D7B0903E1C9D3 CRC64;

Query Match 100.0%; Score 30; DB 11; Length 92;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 HORVH 5
 Db |||||
 21 HORVH 25

 RESULT 14
 O14899 PRELIMINARY; PRT; 102 AA.
 ID O14899
 AC O14899;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Zinc finger protein (ZFP93) (Fragment).
 GN HZF6 OR ZFP93.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98234550; PubMed=9570955;
 RA Shannon M., Stubbs L.;
 RT "Analysis of homologous XRCC1-linked zinc-finger gene families in
 RI human and mouse: evidence for orthologous genes";
 RL Genomics 49:112-121(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kodoyianni V., Ge Y., Krummel G.K., Kvistad E., Grable L.,
 RA Severin J., Gordon L., Shannon M., Brower A., Olsen A.S., Smith L.M.;
 RT "Sequence analysis of a 1mb region in 19q13.2 containing a zinc finger
 RI gene cluster";
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: AF027514; AAD12729.1; -.
 DR EMBL: AC084239; AAG23969.1; -.
 DR HSSP: P08046; 1ALI.
 DR InterPro: IPR007087; Znf_C2H2.
 DR InterPro: IPR007086; Znf_C2H2_sub.
 DR Pfam: PF00096; zf-C2H2; 3.
 DR PRINTS: PR00048; ZINCFINGER.
 DR ProDom: PD000003; Znf_C2H2; 2.
 DR SMART: SM00355; Znf_C2H2; 3.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 4.
 KW Metal-binding; Nuclear protein; Zinc; Zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 102 AA; 11632 MW; 966B8169608186DF CRC64;

 Query Match 100.0%; Score 30; DB 4; Length 102;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 HORVH 5
 Db |||||
 37 HORVH 41

 RESULT 15
 Q9UFH1 PRELIMINARY; PRT; 106 AA.
 ID Q9UFH1
 AC Q9UFH1;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN DKF2P434J0650.
 FT NON_TER 1
 SQ SEQUENCE 106 AA; 11632 MW; 966B8169608186DF CRC64;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Poustka A., Klein M., Meves H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: AL122078; CAB59253.1; -;
 DR InterPro: IPR007087; Znf_C2H2.
 DR InterPro: IPR007086; Znf_C2H2_sub.
 DR Pfam: PF00096; zf-C2H2; 3.
 DR PRINTS: PR00048; ZINGFINGER.
 DR ProDom: PD000003; Znf_C2H2; 2.
 DR SMART: SM00355; Znf_C2H2; 3.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 3.
 KW Hypothetical protein; Metal-binding; Nuclear protein; Zinc;
 KW Zinc-finger.
 FT NON_TER 1
 SQ SEQUENCE 106 AA; 12095 MW; 3D448E0C94F5F16B CRC64;

Query Match 100.0%; Score 30; DB 4; Length 106;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HQRVH 5
 |||||
 Db 44 HQRVH 48

Search completed: October 1, 2003, 10:46:31
 Job time : 31.8611 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:38:06 : Search time 10.4167 Seconds
(without alignments)
20.309 Million cell updates/sec

Title: US-10-057-890A-7

Perfect score: 30

Sequence: 1 HQRVH 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep:*

2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep:*

3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep:*

4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep:*

5: /cgn2.6/ptodata/1/1aa/PCTUS-COMB.pep:*

6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	25	3	US-09-253-396A-216
2	30	100.0	26	2	US-08-620-151-55
3	30	100.0	149	2	US-08-606-143-24
4	30	100.0	150	2	US-08-606-143-26
5	30	100.0	150	2	US-08-606-143-29
6	30	100.0	217	3	US-09-058-489-48
7	30	100.0	288	5	PCT-US95-04801-7
8	30	100.0	315	1	US-08-253-155A-34
9	30	100.0	338	2	US-08-933-750C-4
10	30	100.0	368	3	US-09-234-613-4
11	30	100.0	368	3	US-08-933-750C-14
12	30	100.0	368	3	US-09-234-613-14
13	30	100.0	400	4	US-09-252-991A-25966
14	30	100.0	452	4	US-09-252-991A-20410
15	30	100.0	488	2	US-08-933-750C-17
16	30	100.0	488	3	US-09-234-613-17
17	30	100.0	543	4	US-09-362-123A-9
18	30	100.0	675	1	US-08-317-522A-9
19	30	100.0	675	1	US-08-439-818A-9
20	30	100.0	675	2	US-08-751-965-9
21	30	100.0	675	2	US-08-738-975-9
22	30	100.0	675	2	US-08-728-626-9
23	30	100.0	675	3	US-08-808-599A-9
24	30	100.0	927	4	US-09-252-991A-16765
25	30	100.0	1127	4	US-09-252-991A-26849
26	29	96.7	23	3	US-09-253-396A-218
27	29	96.7	26	2	US-08-620-151-37

28 29 96.7 26 2 US-08-620-151-44 Sequence 44, Appl

29 29 96.7 26 2 US-08-620-151-51 Sequence 51, Appl

30 29 96.7 26 2 US-08-620-151-52 Sequence 52, Appl

31 29 96.7 26 2 US-08-620-151-58 Sequence 58, Appl

32 29 96.7 26 2 US-08-620-151-84 Sequence 84, Appl

33 29 96.7 26 2 US-08-620-151-86 Sequence 86, Appl

34 29 96.7 26 2 US-08-620-151-90 Sequence 90, Appl

35 29 96.7 26 2 US-08-620-151-97 Sequence 97, Appl

36 29 96.7 28 1 US-08-331-394-20 Sequence 20, Appl

37 29 96.7 28 1 US-08-250-858-20 Sequence 20, Appl

38 29 96.7 28 1 US-08-446-915-20 Sequence 20, Appl

39 29 96.7 28 2 US-08-744-139-20 Sequence 20, Appl

40 29 96.7 28 3 US-09-253-396A-225 Sequence 225, App

41 29 96.7 28 4 US-08-779-599-20 Sequence 20, Appl

42 29 96.7 28 5 PCT-US95-06639-20 Sequence 20, Appl

43 29 96.7 41 4 US-09-389-831-15 Sequence 15, Appl

44 29 96.7 41 4 US-09-389-831-16 Sequence 16, Appl

45 29 96.7 60 2 US-08-117-952-785 Sequence 785, App

ALIGNMENTS

RESULT 1

US-09-253-396A-216

: Sequence 216, Application US/09253396A

: Patent No. 6205404

: GENERAL INFORMATION:

: APPLICANT: Genome Dynamics, Inc.

: TITLE OF INVENTION: DNA-Binding Proteins of the Zinc-Finger Class

: FILE REFERENCE: 1116242-0003 file: genome03F.app

: CURRENT APPLICATION NUMBER: US/09/253,396A

: CURRENT FILING DATE: 1999-02-19

: NUMBER OF SEQ ID NOS: 231

: SOFTWARE: Patentin Ver. 2.0

: SEQ ID NO 216

: LENGTH: 25

: TYPE: PRT

: ORGANISM: Xenopus laevis

US-09-253-396A-216

Query Match 100.0%; Score 30; DB 3; Length 25;

Best Local Similarity 100.0%; Pred. No. 5.2;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HQRVH 5

DB 19 HQRVH 23

RESULT 2

US-08-620-151-55

: Sequence 55, Application US/08620151

: Patent No. 5928955

: GENERAL INFORMATION:

: APPLICANT: Imperiali, Barbara

: APPLICANT: Walkup, Grant K.

: TITLE OF INVENTION: PEPTIDYL FLUORESCENT CHEMOSENSOR FOR

: TITLE OF INVENTION: DIVALENT ZINC

: NUMBER OF SEQUENCES: 136

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: BRINKS, HOFER, GILSON & LIONE

: STREET: NBC Tower, Suite 3600, 455 N. Cityfront

: CITY: Chicago

: STATE: Illinois

: COUNTRY: USA

: ZIP: 60611-5599

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,151
; FILING DATE: 22-MAR-1996
; CLASSIFICATION: 422
; ATTORNEY/AGENT INFORMATION:
; NAME: Shannon, Karen L.
; REGISTRATION NUMBER: 36,675
; REFERENCE/DOCKET NUMBER: 8597/6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-620-151-55

Query Match 100.0%; Score 30; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
Db 20 HORVH 24

RESULT 3
US-08-606-143-24
; Sequence 24, Application US/08606143
; Patent No. 5856155
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: COMPOUNDS AND RELATED METHODS FOR
; TITLE OF INVENTION: MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,143
; FILING DATE: 23-FEB-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kilyk Jr., John
; REGISTRATION NUMBER: 30763
; REFERENCE/DOCKET NUMBER: 71756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-606-143-24

Query Match 100.0%; Score 30; DB 2; Length 149;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
Db 1 HORVH 5

RESULT 4
US-08-606-143-26
; Sequence 26, Application US/08606143
; Patent No. 5856155
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: COMPOUNDS AND RELATED METHODS FOR
; TITLE OF INVENTION: MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,143
; FILING DATE: 23-FEB-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kilyk Jr., John
; REGISTRATION NUMBER: 30763
; REFERENCE/DOCKET NUMBER: 71756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-606-143-26

Query Match 100.0%; Score 30; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
Db 1 HORVH 5

RESULT 5
US-08-606-143-29
; Sequence 29, Application US/08606143
; Patent No. 5856155
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: COMPOUNDS AND RELATED METHODS FOR
; TITLE OF INVENTION: MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
```

```
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,143
; FILING DATE: 23-FEB-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kilyk Jr., John
; REGISTRATION NUMBER: 30763
; REFERENCE/DOCKET NUMBER: 71756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-606-143-29

Query Match 100.0%; Score 30; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 1 HORVH 5

RESULT 6
US-09-058-489-48
; Sequence 48, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WH197-08pa
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Human
; US-09-058-489-48

Query Match 100.0%; Score 30; DB 3; Length 217;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 109 HORVH 113

RESULT 7
PCT-US95-04801-7
; Sequence 7, Application PC/TUS9504801
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan F.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enguita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llaarena, Francisco J.
; APPLICANT: Liras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHAMYCIN BIOSYNTHESIS
; TITLE OF INVENTION: LATE GENES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04801
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-04801-7

Query Match 100.0%; Score 30; DB 5; Length 288;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 128 HORVH 132

RESULT 8
US-08-253-155A-34
; Sequence 34, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jeno
; APPLICANT: Draetta, Giulio
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/253.155A
;; FILING DATE: 02-JUN-1994
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Vincent, Matthew P.
;; REGISTRATION NUMBER: 36,709
;; REFERENCE/DOCKET NUMBER: MII-028
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 227-7400
;; TELEFAX: (617) 227-5941
;; INFORMATION FOR SEQ ID NO: 34:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 315 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: internal
US-08-253-155A-34

Query Match 100.0%; Score 30; DB 1; Length 315;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 32 HORVH 36

RESULT 9
US-08-933-750C-4
; Sequence 4, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933.750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 338 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: HMCINOT01
;; CLONE: 9337
US-08-933-750C-4

Query Match 100.0%; Score 30; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
DB 238 HORVH 242

RESULT 10
US-09-234-613-4
; Sequence 4, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234.613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933.750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 338 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 9337
US-09-234-613-4

Query Match 100.0%; Score 30; DB 3; Length 338;
Best Local Similarity 100.0%; Pred. No. 64;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
 Db 238 HORVH 242

RESULT 11
 US-08-933-750C-14
 ; Sequence 14, Application US/08933750C
 ; Patent No. 5932442
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Guegler, Karl J.
 ; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
 ; NUMBER OF SEQUENCES: 98
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/933.750C
 ; FILING DATE: September 23, 1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0356 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 368 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: MMLR3DT01
 ; CLONE: 569710
 ; US-08-933-750C-14

Query Match 100.0%; Score 30; DB 2; Length 368;

Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
 Db 297 HORVH 301

RESULT 12
 US-09-234-613-14
 ; Sequence 14, Application US/09234613
 ; Patent No. 6132973

GENERAL INFORMATION:
 APPLICANT: Lal, Preeti
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Bandman, Olga
 APPLICANT: Shah, Purvi
 APPLICANT: Au-Young, Janice
 APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
 NUMBER OF SEQUENCES: 98
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/234.613
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/933.750
 FILING DATE: September 23, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0356 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 368 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: MMLR3DT01
 CLONE: 569710
 US-09-234-613-14

Query Match 100.0%; Score 30; DB 3; Length 368;

Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
 Db 297 HORVH 301

RESULT 13
 US-09-252-991A-25966
 ; Sequence 25966, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27

```
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25966
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25966

Query Match          100.0%; Score 30; DB 4; Length 400;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HORVH 5
Db      192 HORVH 196
      |||||

RESULT 14
US-09-252-991A-20410
; Sequence 20410, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074.788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094.190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20410
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20410

Query Match          100.0%; Score 30; DB 4; Length 462;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HORVH 5
Db      210 HORVH 214
      |||||

RESULT 15
US-08-933-750C-17
; Sequence 17, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

```
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT03
; CLONE: 641127
; US-08-933-750C-17

Query Match          100.0%; Score 30; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HORVH 5
Db      309 HORVH 313
      |||||

Search completed: October 1, 2003, 10:49:13
Job time : 11.4167 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:42:07 ; Search time 13.3333 Seconds
(without alignments)
59.330 Million cell updates/sec

Title: US-10-057-890A-7
Perfect score: 30
Sequence: 1 HQRVH 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	5	15	US-10-057-890A-7
2	30	100.0	23	12	US-10-223-765-181
3	30	100.0	30	9	US-09-864-761-43845
4	30	100.0	88	10	US-09-764-864-1488
5	30	100.0	93	12	US-10-144-156-2
6	30	100.0	114	10	US-09-764-864-1372
7	30	100.0	115	10	US-09-925-300-1555
8	30	100.0	120	10	US-09-764-864-970
9	30	100.0	125	9	US-09-764-853-815
10	30	100.0	126	9	US-09-864-761-43918
11	30	100.0	127	12	US-10-160-162-238
12	30	100.0	135	15	US-10-106-698-6455
13	30	100.0	138	15	US-10-057-890A-10
14	30	100.0	154	10	US-09-764-864-1321
15	30	100.0	157	15	US-10-057-890A-31

16	30	100.0	165	10	US-09-764-864-1486	Sequence 1486, Ap
17	30	100.0	166	10	US-09-764-864-1472	Sequence 1472, Ap
18	30	100.0	180	9	US-09-864-761-36704	Sequence 36704, A
19	30	100.0	191	10	US-09-764-864-1358	Sequence 1358, Ap
20	30	100.0	192	10	US-09-764-864-1046	Sequence 1046, Ap
21	30	100.0	192	10	US-09-764-864-1473	Sequence 1473, Ap
22	30	100.0	214	9	US-09-864-761-42889	Sequence 42889, A
23	30	100.0	220	15	US-10-106-698-5514	Sequence 5514, Ap
24	30	100.0	223	10	US-09-764-864-973	Sequence 973, Ap
25	30	100.0	254	9	US-09-864-761-33477	Sequence 33477, A
26	30	100.0	257	10	US-09-764-864-1428	Sequence 1428, Ap
27	30	100.0	259	9	US-09-864-761-37325	Sequence 37325, A
28	30	100.0	277	9	US-09-864-761-36043	Sequence 36043, A
29	30	100.0	281	10	US-09-764-864-1076	Sequence 1076, Ap
30	30	100.0	285	9	US-09-764-853-633	Sequence 633, Ap
31	30	100.0	287	9	US-09-864-761-35387	Sequence 35387, A
32	30	100.0	287	10	US-09-764-864-934	Sequence 934, Ap
33	30	100.0	295	10	US-09-764-864-1016	Sequence 1016, Ap
34	30	100.0	301	9	US-09-864-761-47723	Sequence 47723, A
35	30	100.0	303	9	US-09-864-761-36091	Sequence 36091, A
36	30	100.0	313	9	US-09-800-729-196	Sequence 196, Ap
37	30	100.0	322	10	US-09-764-864-914	Sequence 914, Ap
38	30	100.0	326	9	US-09-800-729-195	Sequence 195, Ap
39	30	100.0	333	10	US-09-764-864-958	Sequence 958, Ap
40	30	100.0	333	10	US-09-764-864-1103	Sequence 1103, Ap
41	30	100.0	337	15	US-10-011-585A-178	Sequence 178, Ap
42	30	100.0	338	9	US-09-864-761-46977	Sequence 46977, A
43	30	100.0	338	9	US-09-840-787-4	Sequence 4, Appli
44	30	100.0	342	9	US-09-864-761-36708	Sequence 36708, A
45	30	100.0	368	9	US-09-840-787-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-10-057-890A-7
; Sequence 7, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION: Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 7
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-057-890A-7

Query Match 100.0%; Score 30; DB 15; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HQRVH 5

DB 1 HQRVH 5

RESULT 2
US-10-223-765-181
; Sequence 181, Application US/10223765
; Publication No. US20030165997A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jin-Soo

; APPLICANT: Bae, Kwang-Hee
; APPLICANT: Park, Kyung-Soon
; APPLICANT: Kwon, Young Do
; APPLICANT: Ryu, Eun-Hyun
; APPLICANT: Hwang, Moon-Sun
; TITLE OF INVENTION: ZINC FINGER DOMAIN LIBRARIES
; FILE REFERENCE: 12279-005001
; CURRENT APPLICATION NUMBER: US/10/223,765
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 60/374,355
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 60/313,402
; PRIOR FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 181
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-223-765-181

Query Match 100.0%; Score 30; DB 12; Length 23;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
Db 19 HQRVH 23

RESULT 3

US-09-864-761-43845
; Sequence 43845, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43845
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005324.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.86
; OTHER INFORMATION: SWISSPROT HIT: O60765, EVALUE 3.00e-08
; OTHER INFORMATION: EST_HUMAN HIT: BE902618.1, EVALUE 6.00e-12
US-09-864-761-43845

Query Match 100.0%; Score 30; DB 9; Length 30;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
Db 21 HQRVH 25

RESULT 4

US-09-764-864-1488
; Sequence 1488, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1488
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1488

Query Match 100.0%; Score 30; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
Db 54 HQRVH 58

RESULT 5

US-10-144-156-2
; Sequence 2, Application US/10144156
; Publication No. US20030166197A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R.
; APPLICANT: Nehring, Ramlah
; APPLICANT: McGrath, Robert B.
; TITLE OF INVENTION: ETHYLENE INSENSITIVE PLANTS
; FILE REFERENCE: SALKINS.040A
; CURRENT APPLICATION NUMBER: US/10/144,156


```
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: ZN_FING
; FEATURE:
; LOCATION: (1)...(93)
US-10-144-156-2

Query Match          100.0%; Score 30; DB 12; Length 93;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
Db 23 HQRVH 27

RESULT 6
US-09-764-864-1372
; Sequence 1372, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1372
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1372

Query Match          100.0%; Score 30; DB 10; Length 114;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
Db 15 HQRVH 19

RESULT-7-
US-09-925-300-1555
; Sequence 1555, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PRT101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1555
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (51)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1555

Query Match          100.0%; Score 30; DB 10; Length 115;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
Db 62 HQRVH 66

RESULT 8
US-09-764-864-970
; Sequence 970, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 970
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-970

Query Match          100.0%; Score 30; DB 10; Length 120;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
Db 115 HQRVH 119

RESULT 9
US-09-764-853-815
; Sequence 815, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRT06
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 815
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-815

Query Match          100.0%; Score 30; DB 9; Length 125;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HQRVH 5
Db 74 HQRVH 78
```

RESULT 10

US-09-864-761-43918
; Sequence 43918, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 4917
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43918
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006213.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.97
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: Q9UK11, EVALUATE 2.00e-50
US-09-864-761-43918

Query Match 100.0%; Score 30; DB 9; Length 126;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5

Db 115 HORVH 119

RESULT 11

US-10-160-162-238
; Sequence 238, Application US/10160162
; Publication No. US20030166541A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 83 Human Secreted Proteins
; FILE REFERENCE: P2012P2
; CURRENT APPLICATION NUMBER: US/10/160,162
; CURRENT FILING DATE: 2002-06-04
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 60/295,558
; PRIOR FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: 09/236,557
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: PCT/US98/15949
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,212
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,209
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,234
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,218
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,214
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,236
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,215
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,211
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,217
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,213
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/055,968
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,969
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,972
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,561
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,534
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,729
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,543
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,727
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,554
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,730
; PRIOR FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 238
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-160-162-238

Query Match

100.0%; Score 30; DB 12; Length 127;

Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
 |||||
Db 46 HORVH 50

RESULT 12
US-10-106-698-6455
: Sequence 6455, Application US/10106698
: Publication No. US20030109690A1
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
: FILE REFERENCE: PA005P1
: CURRENT APPLICATION NUMBER: US/10/106,698
: CURRENT FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: PCT/US00/26524
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: US 60/157,137
: PRIOR FILING DATE: 1999-09-29
: PRIOR APPLICATION NUMBER: US 60/163,280
: PRIOR FILING DATE: 1999-11-03
: NUMBER OF SEQ ID NOS: 8564
: SOFTWARE: Patentin Ver. 3.0
: SEQ ID NO 6455
: LENGTH: 135
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (11)
: NAME/KEY: MISC_FEATURE
: LOCATION: (12)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (13)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (24)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (50)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (64)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (66)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (69)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (70)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (98)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (102)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (118)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (126)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (131)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-10-106-698-6455

Query Match 100.0%; Score 30; DB 15; Length 135;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
 |||||
Db 110 HORVH 114

RESULT 13
US-10-057-890A-10
: Sequence 10, Application US/10057890A
: Publication No. US20030044901A1
: GENERAL INFORMATION:
: APPLICANT: Coleman, Timothy
: TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same,
: FILE REFERENCE: PF537
: CURRENT APPLICATION NUMBER: US/10/057,890A
: CURRENT FILING DATE: 2002-01-29
: PRIOR APPLICATION NUMBER: 60/265,782
: PRIOR FILING DATE: 2001-01-31
: PRIOR APPLICATION NUMBER: 60/265,858
: PRIOR FILING DATE: 2001-01-31
: NUMBER OF SEQ ID NOS: 32
: SEQ ID NO 10
: LENGTH: 138
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-057-890A-10

Query Match 100.0%; Score 30; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HORVH 5
 |||||
Db 50 HORVH 54

RESULT 14
US-09-764-864-1321
: Sequence 1321, Application US/09764864
: Patent No. US20020132753A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT223
: CURRENT APPLICATION NUMBER: US/09/764,864
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1792
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1321
: LENGTH: 154
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (103)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (112)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1321

Query Match 100.0%; Score 30; DB 10; Length 154;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
 Db 14 HORVH 18

RESULT 15
 US-10-057-890A-31
 ; Sequence 31, Application US/10057890A
 ; Publication No. US20030044901A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Coleman, Timothy
 ; APPLICANT: Mansfield, Brian
 ; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, and
 ; TITLE OF INVENTION: of Using the Same.
 ; FILE REFERENCE: PF537
 ; CURRENT APPLICATION NUMBER: US/10/057,890A
 ; CURRENT FILING DATE: 2002-01-29
 ; PRIOR APPLICATION NUMBER: 60/265,782
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/265,858
 ; PRIOR FILING DATE: 2001-01-31
 ; NUMBER OF SEQ ID NOS: 32
 ; SEQ ID NO 31
 ; LENGTH: 157
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-057-890A-31

Query Match 100.0%; Score 30; DB 15; Length 157;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HORVH 5
 Db 69 HORVH 73

Search completed: October 1, 2003, 10:50:54
 Job time : 14.3333 secs